Arthaniti

(New Series)

Volume ~~IV~~ V, Nos. (1-2), ~~2005~~ 2005

EDITORIAL BOARD

Chairman
Soumyen Sikdar

Joint Editors
Sarbajit Chaudhuri
Sankar Kumar Bhaumik

Associate Editors
Ishita Mukhopadhyay
Rabindranath Mukhopadhyay

Book Review Editors
Nita Mitra
Anjan Chakraborty

Editorial Secretary
Kanai Lal Basak

Members
Asis Kumar Banerjee
Arup Mallick
Ajit Chaudhuri
Kalyan Sanyal
Sarmila Banerjee
Pabitra Kumar Giri
Manash Ranjan Gupta
Sugata Marjit
Dipankar Coondoo
Abhirup Sarkar

Editorial Advisory Board
Ronald W Jones
Timothy Besley
Sajal Lahiri
Partha Sen
Prasanta K. Pattanaik
Rabindra Nath Bhattacharya
Prabuddha Nath Roy
Amiya K Bagchi
Dilip Mookherjee
Pasquale Sgro
Debraj Ray
Hiron Kumar Banerjee
David E A Giles
Bharat R Hazari
Hamid Beladi
Graciela Chichilnisky
Ira Gang
Shigemi Yabuuchi
Ulaganathan Sankar

Contents

How workers get poor because capitalists get rich : A general equilibrium model of labor supply, community and the class distribution of income
Indraneel Dasgupta and Rabi Kanbur

On the incentives for cooperative and non-cooperative R&D in duopoly
Tarun Kabiraj

Informal enterprises : Reflections on existing hypotheses
Vinod Anand

International trade and personal income distribution in developing countries : Some issues and concerns
Ananya G. Dastidar

Dirtier trade for India? The story of globalization
Sudeshna Chattopadhyay

Monetary models of exchange rates and the random walk
Gianluca Lagana and Pasquale M Sgro

CENTRAL LIBRARY
UNIVERSITY

EDITORIAL/ BUSINESS OFFICE
Department of Economics
University of Calcutta
56A, B.T. Road, Kolkata 700 050, India.
Tel: 91-33-2557-5082, 91-33-2528-4066
Fax: 91-33-2528-4066
E-mail: sarbajitch@yahoo.com
spsikdar@cal3.vsnl.net.in

⑤

ARTHANITI

(NEW SERIES)

THE JOURNAL OF THE DEPARTMENT OF ECONOMICS, UNIVERSITY OF CALCUTTA

Aims & scope: The Dept. of Economics, University of Calcutta is resuming the publication of its journal *ARTHANITI* in 2002 as a new series. For the time being, *ARTHANITI* will be published as an annual international journal. Papers across the entire field of economics, including econometrics, economic history, and political economy will be considered for publication. Both theoretical and empirical manuscripts are equally welcomed.

Manuscripts will be judged solely on the quality of their content. *ARTHANITI* will adopt a double-blind review process. The editorial board will not reveal the identities of either authors or referees to other parties in the review process. The editor and/or any member of the editorial board will send submitted papers to the referees only after an initial reading. The editorial board reserves the right to reject any paper at that stage. However, most of the submitted papers will be sent to the referees for comments.

SUBSCRIPTION RATES:

G165642

Indian: Institutions:	Rs. 300/- (Annual)	Overseas: Institutions:	US $ 40 (Annual)
	Rs. 750/- (3 Years)		US $ 100 (3 Years)
Individuals:	Rs. 150/- (Annual)	Individuals:	US $ 25 (Annual)
	Rs. 375/- (3 Years)		US $ 60 (3 Years)
		Air Surcharge:	US $ 5 (optional)

without the written consent of the publisher.

Published by Soumyen Sikdar, Chairman, Editorial Board of ARTHANITI, on behalf of the Department of Economics, Calcutta University.

Printed at: Calcutta University Press

Computer Type Setting & Designing : Suchhanda Chaudhury

Front Page Design : Suchhanda Chaudhury

Cuk H00555-5-U2165642

135/5

ARTHANITI

New Series

Volume V, Nos. (1-2), 2006

Contents

How workers get poor because capitalists get rich : A general equilibrium model of labor supply, community and the class distribution of income

Indraneel Dasgupta and Rabi Kanbur 1-23

On the incentives for cooperative and non-cooperative R&D in duopoly

Tarun Kabiraj 24-33

Informal enterprises : Reflections on existing hypotheses

Vinod Anand 34-63

International trade and personal income distribution in developing countries : Some issues and concerns

Ananya G. Dastidar 64-80

Dirtier trade for India? The story of globalization

Sudeshna Chattopadhyay 81-95

Monetary models of exchange rates and the random walk

Gianluca Lagana and Pasquale M Sgro 96-113

Arthaniti 5 (1-2), 1-23 / 2006

How Workers Get Poor *Because* Capitalists Get Rich: A General Equilibrium Model of Labor Supply, Community, and the Class Distribution of Income*

Indraneel Dasgupta[A] and Ravi Kanbur[B]

Abstract

We develop an integrated, general equilibrium, model of how the presence of vertical ties of 'community' between sections of workers and sections of capitalists can critically affect the distribution of income between capitalists as a class and workers as a class, as well as between workers belonging to different communities. We show that an exogenous increase in the incomes of capitalists sets in motion community and market processes that subsequently (a) further increase capitalists' incomes, (b) can reduce workers' earnings as well as welfare, and (c) systematically influence earnings differentials between workers belonging to different communities.

Keywords : Community, Class, Public Goods, Income Distribution

JEL Classification : D31, D63, D74, Z13.

---

1. Introduction

Inequality in earnings has increased significantly in many countries over the last two decades or so.[1] One aspect of this is the increase in the earnings gap between workers and those who own capital assets. Another aspect is the increase in the earnings gap between workers belonging to different religious, ethnic or language groups.[2] Both aspects have been documented and debated at length. Our paper contributes to this literature by providing one explanation of how an initial increase in the incomes of capital owners, regardless of its cause, can feed back into the economy to: (a) reduce the earnings of workers of all communities, (b) increase the earnings gap between workers belonging to different communities even when there is no segmentation or discrimination in the labor market, and (c) increase the incomes of capital owners even further.

---

* We thank Prasanta Pattanaik for helpful discussions, and The Pew Charitable Trusts for financial support.

A University of Nottingham, indraneel.dasgupta@nottingham.ac.uk

B Cornell University, sk145@cornell.edu

[1] For recent surveys, see Kanbur (2000) and Kanbur and Lustig (2000).

[2] There are numerous analyses of ethnic inequality in the US. An early attempt at rigorous analysis of inter-ethnic inequality in a developing country context was by Anand (1983).

Two strands of thought motivate our analysis—the first emanating from a social observation, and the second from a question. The social observation is that "vertical" ties of community cut across "horizontal" class difference between poor and rich individuals. These community ties can be of different types—ethnic, religious, clan, etc—but they exert a pull and affect behavior over and above class position. This is the observation. The question is as follows: can the poor become poorer *because* the rich become richer? Or, to put the question differently, if the rich become richer for whatever reason, could this event, simply by itself, generate forces that would subsequently reduce the welfare of the poor? Both the observation and the question need elaboration. We start with the question.

The claim that prior wealth gains by the rich can somehow be *causally responsible* for subsequently *aggravating* poverty may appear perplexing in the light of standard economic theory. Typically, rich individuals are those who own significant amounts of non-labor factors of production, in particular, land or physical capital, and derive their incomes mainly from such ownership. Thus, by and large, rich individuals are *capitalists*, whether industrial or agrarian. Conversely, poor individuals are mostly *workers*, i.e., those who do not own significant amounts of non-labor assets, and are primarily dependent on the labor market for their incomes. Wealth gains by capitalists would imply either (a) returns from their existing holdings of land or capital assets have gone up for some reason, or (b) they have somehow come into possession of more land or capital stock, or both. Reasons typically cited for possibility (a) include population growth, biased technological progress, state subsidies funded by additional taxes on workers, and demand driven shifts in relative prices in favor of land-intensive or capital-intensive goods, which reduce the wage rates, and thus, the earnings of the workers. However, in this case, it is some extraneous demographic, technological, politico-economic or demand side factor that is causally responsible for *both* income losses by workers and income gains by capitalists: the latter does not cause the former. Thus, workers become poorer *while*, but *not because*, capitalists become richer.[3] If possibility (b) obtains, then, since total land or capital stock in the economy rises, labor demand should increase, increasing wage rates and/or employment, thereby making workers *better off*. Thus, claims regarding the existence of causal mechanisms connecting prior income gains by the rich (capitalists) with subsequent income losses by the poor (workers) are often discounted as dubious. We attempt to rehabilitate such claims by advancing one causal mechanism in this paper.

---

[3] It is in this sense that Proudhon's oft-quoted dictum about property being theft under capitalism is a condemnation of wealth ownership on moral, but not functional grounds. In his argument, it is the legal framework of private ownership of capital which allows capitalists to acquire wealth which 'rightfully' belongs to workers, and thereby impoverishes workers, but such wealth, once acquired, does not generate *further* impoverishment of the workers. See Proudhon (1994).

In standard analyses of the class basis of income distribution, and of distributive conflicts between capitalists and workers,[4] the two classes are linked with one another only through the market (employment) relationship.[5] The argument we develop builds however on the observation that in many, if not most, societies some sections of capitalists and some sections of workers are also vertically linked through social ties of 'community', whereas other sections of workers are excluded from this relationship, by virtue of their belonging to a different community.

Most ethnic, religious, racial, caste or geographic communities are characterized by internal differentiation and heterogeneity with respect to ownership of productive non-labor assets, and, consequently, income levels. At the same time, what seems to define such communities as communities is the presence of something, (a) to which all members have common access, but from which non-members are excluded, and (b) in the process of accessing which members also acquire social access to each other, despite differences in their income and wealth status. Thus, one way of formalizing the notion of *belonging* to a community is by positing access to a community-specific local public good.[6] For example, belonging to a particular religious community can be interpreted to imply (a) common access to that community's places of worship and religious activities, and (b) access to social interaction with other members of the community *via* common participation in worship and collective religious activities. Similarly, the right to participate in ethno-linguistic festivals and rituals can be considered constitutive of the ethnic identity of an individual, and participation in such activities may provide a platform, perhaps the only platform, where rich and poor members of the ethnic group come together to interact socially.[7] At a slightly different level, the presence of local or neighborhood public goods (such as parks, roads, libraries, concert halls, auditoria, museums, sports clubs, safety etc.), and the social interaction between rich and poor which is brought about as an indirect consequence of accessing and using such public goods, can be considered as defining that which constitutes a *community* out of a collection of individuals living in geographical proximity to each other.

---

[4] See, for example, Lancaster (1973), Przeworski and Wallerstein (1982), Alesina and Rodrik (1994) and Somanathan (2001).

[5] An exception is Roemer (1998), who explicitly considers the role played by ties of religion in the determination of the level of redistribution in a political-economic equilibrium.

[6] Such a public (or rather, 'club') good-based interpretation of community identity and membership is developed in detail by Dasgupta and Kanbur (2001).

[7] There is a large anthropological literature on community identity. Economists have recently become interested in this notion. See Akerlof and Kranton (2000). The link of public goods provision and ethnic divisions has been explored in Alesina, Baqir and Easterly (1999). There is also a psychological literature on group identity, as exemplified by Wetherell (1996).

The productivity of a worker can be expected to increase with the size of the public good(s) which defines the community to which he belongs. This is for direct as well as indirect reasons. Most obviously, better infrastructure, in the form of a larger quantum, or improved quality, of local public goods of the kind mentioned above, can be expected to improve her physical and psychological health, analytical abilities, knowledge and awareness of the world, reduce the stress and exertion related to commuting etc., all of which would improve her ability to work more intensely, or allow her to maintain high levels of efficiency for longer periods. An expansion in the magnitude of the community specific public good would also expand the scope for social interaction with other poor individuals, as well as rich individuals. This can lead to more information pooling, improvements in inter-personal social skills, learning of better work, health and recreation practices, finding job environments which provide better fits for one's individual idiosyncratic characteristics and abilities, etc.[8] Thus, workers' productivity may go up because of effects working through this indirect social channel as well.[9]

The intuitive form of the argument we develop and formalize is the following. Consider a society comprised of two, say ethnic, communities, A and B. Suppose that, for some reason, capitalists belonging to community A acquire more wealth. They may be expected to spend a part of their additional wealth on public goods specific to their community. This, however, for reasons already discussed, will improve the productivity of workers of community A. Consequently, labor supply of workers of community A, measured in efficiency units, will go up, reducing the wage rate per efficiency unit of labor supply. Additional labor supply, will, however, increase returns to capital ownership earned by all capitalists (irrespective of community affiliation). This second-order effect, while improving the welfare of every capitalist in society, will also increase labor supplies of both communities, thereby further depressing the wage rate. The overall fall in the wage rate will reduce the welfare of *every* worker in society, if aggregate labor demand is sufficiently inelastic. Indeed, even if labor demand is elastic, workers from one community can suffer a reduction in their welfare if their productivity gains are much lower than those of workers from the other community. Regardless of the elasticity of labor demand, the earnings gap between workers belonging to different communities can go up with an exogenous increase in the incomes of capitalists, simply because the effect of such an increase on the productivity of workers in community A can be different from those in community B.

---

[8] These "social network" benefits to individuals are well discussed and documented. See, for example, Loury (2002), Ligon, Thomas and Worrall (2002) and Fachamps and Minten (1995).

[9] That socialization with people from diverse economic backgrounds, especially in one's youth, may have a positive impact on one's own productivity is an idea that underlies, at least in part, deliberate attempts by many governments to encourage class heterogeneity within student populations in public educational institutions; and to subsidize housing for the poor in affluent cities and neighborhoods. The understanding that individuals from certain communities, particularly those born into ethnic, religious or racial minority groups, may have less opportunity for such beneficial social interaction within civil society often drives measures to make state institutions more 'inclusive', in terms of ethnic, religious, racial or caste composition, in their recruitment practices. The same understanding, somewhat paradoxically, often also motivates and justifies coercive state measures designed to force minority groups to 'integrate', i.e. adopt ethnic, religious, caste, cultural or language practices of the majority group.

The outline of the paper is as follows. Section 2 sets up the basic framework. Section 3 establishes the general equilibrium in the model. Section 4 conducts comparative statics with respect to an exogenous increase in capitalists' incomes and establishes the basic proposition of the paper. Section 5 concludes.

2. The framework

The picture we wish to paint is that of a society of communities. Each community consists of two classes of agents: capitalists (those who own capital stock) and workers (those who own only their labor power). Production takes place according to a constant returns to scale aggregate production function, labor and capital markets are competitive, and there is no discrimination on the basis of community affiliation in the labor market. What defines community is access to a public good. All agents consume a private good, a community specific public good, and leisure. They face a common wage rate per efficiency unit in the labor market, and are endowed with one time unit of labor. The number of efficiency units of labor that are generated from one time unit, however, depends on, and increases with, the amount of the public good specific to the community to which the agent belongs. All agents choose their labor supply and allocation of their total income between the private good and the public good of their own community in Cournot fashion.

For a formalization of this picture, consider a society consisting of n individuals, $n \geq 4$. The set of individuals is $N = \{1,...,n\}$. There is an exogenously given amount of capital (or land) stock in the economy, K. The society contains two classes, rich (R) and poor (P). The rich are capitalists (or landlords): ownership of the total capital stock, K, is shared equally among all R agents in society. The poor are workers: no P agent owns any capital. The numbers of rich and poor individuals are, respectively, n_R and n_P. Each class contains at least 2 agents, and, obviously, $n_R + n_P = n$.

Society is divided into two communities, A and B. Each community consists of some R individuals, and some P individuals. The number of P individuals belonging to community A is n_{PA}, while the number of P individuals belonging to community B is n_{PB}. Thus, $n_P = n_{PA} + n_{PB}$. Similarly, $n_{RA} + n_{RB} = n_R$.[10]

---

[10] We can allow capital ownership of R individuals to vary across communities without affecting the results.

We first model the production side of our economy.

A. Production and labor demand:

Our society produces a single good, by means of competitive capital and labor markets, according to the constant returns to scale aggregate production function:

$$Q = Q(K, \lambda),$$

where K is the total (exogenously given) amount of capital stock in the economy, and λ is the total amount of labor supplied, measured in efficiency units (to be determined endogenously as an equilibrium outcome); $Q_K, Q_\lambda > 0, Q_{KK}, Q_{\lambda\lambda} < 0, Q_{K\lambda} > 0$, $Q_{K\lambda\lambda} \leq 0$. Thus, if $\lambda^*|$ is the amount of labor supplied in equilibrium, then each R agent earns the amount $\left[\frac{K}{n_R} Q_K\left(K, \lambda^*\right)\right]$ as rental income, while the wage rate per efficiency unit of labor is: $\left[Q_\lambda\left(K, \lambda^*\right)\right]$. The inverse demand function for labor is given by:

$$w = d\left(K, \lambda^D\right) \equiv Q_\lambda\left(K, \lambda^D\right),$$

so that, given capital stock K, if firms employ λ^D efficiency units of labor, the wage prevailing in the labor market must be $d\left(K, \lambda^D\right)$, $d_{\lambda^D} < 0$, $d_K > 0$. We assume that the wage function is bounded from above by some $\overline{w}$, i.e., $Q_\lambda\left(K, 0\right) = \overline{w}$, where $\overline{w} \in \Re_{++}$.

Our primary interest lies in isolating and investigating the consequences of an increase in earnings of R individuals, irrespective of the origins of this increase. We thus assume that, in addition to her rental income, each R agent belonging to community k has some non-negative endowment of the single good in our model, D_k, so that her total income is $\left[D_k + \frac{K}{n_R} Q_K\left(K, \lambda\right)\right]$. To fix ideas, one can think of the economy receiving an amount $\left[n_{RA} D_A + n_{RB} D_B\right]$ of the good from outside the country as foreign inflow, or "manna from heaven", which is divided between R agents of the two communities in some fashion, with all R agents within a community receiving identical amounts. No poor agent in society receives such an exogenous payment.[11] We shall analyze the comparative static properties of our model by looking at its equilibrium response to increases in D_k.

---

[11] This is only for convenience of exposition. We can generalize the treatment by allowing the poor to receive some foreign inflow as well.

All agents live for one period, and all capital stock is used up in the process of production, so that total consumption is simply total output plus total foreign inflows. We thus abstract from dynamic investment considerations.

B. Consumption and labor supply:

Each individual consumes leisure, a private good and a public good that is specific to the community she belongs to, and is endowed with one unit of labor time. Thus, individuals purchase the single commodity that is produced in the economy, and can put it to alternative uses: some uses are characterized by rivalry in consumption, other methods of use entail non-rivalrous consumption within the community. For any individual $i \in N$, belonging to community *k*, $k \in \{A, B\}$, amounts of leisure, the private good and the public good consumed are, respectively, l_i, x_i and y^k. Thus, individuals belonging to community A only have (costless) access[12] to the public good specific to their community, and similarly for individuals belonging to community B.[13]

For an agent *i*, belonging to community *k*, preferences are given by:

$$u = g\left(x_i, y^k\right) - L_i,$$

where $L_i = \left(1 - l_i\right)$ is the amount of labor, in time units, supplied by the agent, *g* is increasing, twice continuously differentiable and strictly quasi-concave in its arguments.[14]

Our first major assumption regarding preferences is the following.

A1. (i) For all $y_k \in \Re_{++}$, $\underset{x_i \to 0}{Lt} \dfrac{g_x\left(x_i, y^k\right)}{g_y\left(x_i, y^k\right)} = \infty$, and for all $x_i \in \Re_{++}$,

---

[12] The assumption of costless access for community members is for notational simplicity. We can allow the possibility that community members have to pay a fixed amount, a 'membership fee', without changing our results.

[13] The public goods specific to the two communities can be considered identical in character, as, for example, might be the case with civic amenities in two cities, but residents of one city would choose not to access the amenities of another city because travel or relocation costs are two high. Alternatively, in cases of, say, religious or ethnic communities, individuals born into community A would not derive any utility from the public good, say temples, definitive of the other community.

[14] The assumption of utility functions being identical across communities is for notational simplicity. We can work with the more general formulation that, for an individual i belonging to community k, preferences are given by $g^k\left(x_i, y^k\right) - \rho_k L_i$, with $\rho_k > 0$, thereby allowing preferences to vary across communities, without affecting our results.

$$\underset{y^k \to 0}{Lt} \frac{g_x(x_i, y^k)}{g_y(x_i, y^k)} = 0 .$$

(ii) $g_{xx}, g_{yy} < 0, g_{xy} \geq 0$.

Remark 2.1. A1(i) ensures that all agents must consume positive amounts of both private and public goods in an equilibrium. A1(ii) implies $\frac{g_x(x_i, y^k)}{g_y(x_i, y^k)}$ is increasing in y^k and decreasing in x_i. Together, A1(i) and A2(ii) ensure that, for an agent with the utility function g, both the private and public goods are normal goods in the standard sense.

The productivity of an hour's labor provided by an individual depends on the amount of the public good that she can consume. More specifically, when an agent *i*, belonging to community *k*, provides L_i amount of market labor in time units, the amount of market labor provided by her in efficiency units is given by:

$$e_i = f(y^k)L_i , \tag{2.1}$$

where *f* is increasing, twice continuously differentiable and strictly concave, with $[f(0) = 1]$, $\left[f'(0) \in \left(0, \frac{1}{\overline{w}}\right)\right]$ and $\left[\underset{y^k \to \infty}{Lt} f(y^k) = \delta\right]$, such that $\delta \in (1, \infty)$. Thus, no agent can have a labor endowment in efficiency units more than δ, which is finite and greater than unity. The maximum possible marginal gain in labor efficiency, $f'(0)$, is also finite, and such marginal gain decreases with an increase in the community's stock of the public good. It follows that society's labor supply can never be more than δn. Define $\underline{w} = d(n\delta)$. Thus, the wage rate in the labor market must lie within the interval $[\underline{w}, \overline{w}]$. We shall assume that $\underline{w} > 0$.

Agents are differentiated, because of their community affiliations, in terms of their consumption of the (community-specific) public good. Consequently they may, possibly, vary in terms of their endowments of labor, when measured in efficiency units, even though each of them owns one time unit of labor. However, agents are integrated in the labor market, in that, every agent in society, regardless of community affiliation, faces a wage rate of w per *efficiency* unit in the labor market, $w \in [\underline{w}, \overline{w}]$.[15]

---

[15] In other words, while community identity and membership for individuals is not an act of choice, but something given, by birth, upbringing or some other historical factor, competitive employers do not discriminate on the basis of community identity in their hiring practices.

Thus, the wage rate per *time* unit of labor, faced by a member of community k, is simply $wf(y^k)$. Let the non-labor income of an individual belonging to class j and community k be I_{jk}. Since P agents do not own any capital stock, nor do they receive any foreign flows, $I_{PA} = I_{PB} = 0$; while $I_{RA}, I_{RB} \geq 0$.

Individuals take the market wage rate per efficiency unit, w, as given, and simultaneously choose their labor supply, as well as the allocation of their total incomes between the private good and the public goods of their respective communities.

Consider an agent i who belongs to class j and community k. Given any amount of the public good contributed by all other members of her community, y_{-i}^k, her own optimal consumption bundle is given by the solution to:

$$\underset{x_i, y^k, L_i}{Max} \; g(x_i, y^k) - L_i \tag{P1}$$

subject to the budget constraint:

$$x_i + y^k = I_{jk} + wf(y^k)L_i + y_{-i}^k, \tag{2.2}$$

and the additional constraint:

$$y^k \geq y_{-i}^k. \tag{2.3}$$

The net price effectively paid by an agent belonging to community k, for an additional unit of the public good of her community, is thus $[1 - wf'(y^k)L_i]$. Since $L_i \in [0,1]$, and, by assumption, $\left[f'(0) \in \left(0, \frac{1}{\overline{w}}\right)\right]$ and $f'' < 0$, this net price is always positive, but not more than 1. Let the minimum net price possible be $\underline{p}$. Clearly, $\underline{p} = (1 - \overline{w}f'(0))$, which implies $\underline{p} \in (0,1)$.

3. Equilibrium

A general equilibrium in our model must have three features. First, the labor market must be in equilibrium. Second, the capital market must be in equilibrium. Third, given the equilibrium wage rate, say w^*, and capitalists' equilibrium rental incomes $I_{Rk}^* - D_k = \frac{K}{n_R} Q_K(K, d^{-1}(w^*))$, $k \in \{A, B\}$, the total amount of labor supply in efficiency units generated by community k, in the Nash equilibrium of the game defined by (P1), (2.2) and (2.3) and corresponding to $\langle I_{Rk}^{*}, w^{*} \rangle$, say L_k^{*}, must be such that $[L_A^{*} + L_B^{*} = d^{-1}(w^{*})]$. In this section, we address the questions of existence, uniqueness and characteristics of such a general equilibrium.

In Section 1, we identified the poor with workers. Our first step is to ensure that, in equilibrium, P individuals, regardless of their community affiliation, all turn out to be agents who provide market labor. To this end, we assume the following.

A2. $g_x(\overline{w}\delta, 0) \geq \frac{1}{\underline{w}}$.

The following implication is then immediate.

Lemma 3.1. *Given A1 and A2, and given any* $w \in [\underline{w}, \overline{w}]$, *it must be the case that, for every P individual in society, the amount of labor, in time units, supplied in any Nash equilibrium corresponding to w is 1.*

Thus, all P agents, regardless of their community affiliations, will necessarily sell their entire endowments of labor power in any Nash equilibrium.[16]

Remark 3.2. It follows from Lemma 3.1 that, given A1 and A2, the amount of labor supplied in efficiency units in any labor market equilibrium must be at least n_P. Hence, if both capital and labor markets are in equilibrium, then, for any R agent of community k, $I_{Rk} - D_k \geq \left(\frac{K}{n_R}\right) Q_K(K, n_P)$.

We now take up the issue of labor supply by R agents. As mentioned in Section 1, by rich individuals we understand individuals whose primary source of income is non-labor earnings. In line with this characterization, we wish to ensure that, in equilibrium, R individuals do in fact turn out to be the idle rich, i.e., those who do not participate in the labor market, preferring, instead, to live off the rental income from their asset holdings, and foreign flows.

Consider the problem:

$\underset{x_l, y^k}{Max}\, g(x_l, y^k)$ subject to:

$x_l + py^k = r$.

Let $V(r, p)$ be the indirect utility function corresponding to g, derived by solving this problem, in the standard way. Obviously, $V_r(r,p)>0$. We assume the following.

---

[16] Allowing P agents to sell only part of their labor endowment complicates the exposition, but does not add anything by way of insight.

A3. (i) For every $p\in[\underline{p},1]$, (a) $V_{rr}(r,p)<0$, and (b) $\underset{r\to\infty}{Lt}\, V_r(r,p)=0, \underset{r\to 0}{Lt}\, V_r(r,p)=\infty$.

(ii) For every $k\in\{A,B\}$,

$$\left(\frac{K}{n_R}\right)Q_K(K,n_P)+D_k > max\{r\in\Re \mid \overline{w}\delta V_r(r,p)=1 \text{ for some } p\in[\underline{p},1]\}.$$

Assumptions A1, A2 and A3, together, will suffice to ensure that, in equilibrium, our R agents will indeed turn out to be the idle rich, or rentier-capitalist coupon clippers.

Lemma 3.3. *Given A1 and A3,* $k\in\{A,B\}$, *any* $w\in[\underline{w},\overline{w}]$, *any* $n_{Pk}\geq 0$, *and any* $I_{Rk}\geq D_k+\left[\frac{K}{n_R}(Q_K(K,n_P))\right]$, *no R agent in community k will provide market labor in any Nash equilibrium corresponding to* $\langle I_{Rk},w\rangle$.

Proof: See the Appendix.

We now proceed to address the question of contributions to the public goods. Consider a counterfactual community consisting only of R individuals belonging to community k, n_{Rk} in number, and suppose (a) $w\in[\underline{w},\overline{w}]$, and (b) $I_{Rk}\geq D_k+\left[\frac{K}{n_R}(Q_K(K,n_P))\right]$. By Lemma 3.3, these individuals will consume their entire endowments of leisure, and thus, spend only their non-labor earnings, I_{Rk} each, on the private and public goods. Let the amount of the public good of community k generated in the corresponding Nash equilibrium be given by $\alpha(I_{Rk},n_{Rk})$.[17] It is easy to check that the function α is increasing in its arguments. We shall assume that $\alpha_{11}\leq 0$.

In light of Lemma 3.3, it then follows that, given A1 and A3, the Nash equilibrium level of the public good in our original, two-class, community k must be at least $\alpha(I_{Rk},n_{Rk})$. It will be exactly this amount if only R agents are contributory; it will be more than $\alpha(I_{Rk},n_{Rk})$ if P agents are contributory as well. Our next step is to rule out the latter case.

---

[17] A1 suffices to guarantee the existence and uniqueness of such a Nash equilibrium. See Remark 2.1 and Bergstrom, Blume and Varian (1986).

A4. $\underset{y^k \to \infty}{Lt} \dfrac{g_x\left(\overline{w}\delta, y^k\right)}{g_y\left(\overline{w}\delta, y^k\right)} = \infty$.

A1 and A4 together yield:

there exists $\widetilde{y} \in \Re_{++}$ such that: $\left[\dfrac{g_x(\overline{w}\delta, \widetilde{y})}{g_y(\overline{w}\delta, \widetilde{y})} = \dfrac{1}{1-\overline{w}\delta}\right]$.

Given A1(ii) (Remark 2.1), and any $w \in \left[\underline{w}, \overline{w}\right]$, it immediately follows that:

$$\text{for all } y^k \geq \widetilde{y}, \left[\frac{g_x\left(wf\left(y^k\right), y^k\right)}{g_y\left(wf\left(y^k\right), y^k\right)} > \frac{1}{1-wf\left(y^k\right)}\right]. \tag{3.1}$$

(3.1) essentially implies that, if capitalists of community k spend a sufficiently large amount on the public good defining their community, then workers of their community will spend nothing on the public good, preferring, instead, to completely free-ride on the capitalists. Our next assumption is simply that capitalists indeed spend such a large amount.

A5. For all $k \in \{A, B\}$, $\alpha\left(D_k + \dfrac{K}{n_R}\left(Q_K\left(K, n_P\right)\right), n_{Rk}\right) \geq \widetilde{y}$.

Noting that A1 and Lemma 3.3 together imply that, given $I_{Rk} \geq D_k + \left[\dfrac{K}{n_R}\left(Q_K\left(K, n_P\right)\right)\right]$, in a Nash equilibrium, $y^k \geq \alpha\left(I_{Rk}, n_{Rk}\right)$, and using (3.1), we have the following result.

Lemma 3.4. *Given A1-A5, any* $k \in \{A, B\}$, *any* $I_{Rk} \geq D_k + \left[\dfrac{K}{n_R}\left(Q_K\left(K, n_P\right)\right)\right]$, *and any* $w \in \left[\underline{w}, \overline{w}\right]$, *in every Nash equilibrium corresponding to* $\left\langle I_{Rk}, w\right\rangle$,

(i) no P agent in k will contribute to her community's public good,

and (ii) $y^k = \alpha(I_{Rk}, n_{Rk})$.

Remark 3.5. In light of the preceding discussion, it is easy to see that, given A1-A5, and given any $w \in [\underline{w}, \overline{w}]$, and any $I_{Rk} \geq D_k + \left[\frac{K}{n_R}(Q_K(K, n_P))\right]$, both existence and uniqueness of the Nash equilibrium must be satisfied for community k.

In our model, rental incomes of capitalists of a community determine the level of public good in that community. The latter, in turn, determines the productivity of workers in the community, and thus, aggregate labor supply (in efficiency units) in the society. Total labor supply, in turn, determines the marginal product of capital, and thereby, rental incomes of capitalists. We now turn to the problem of equilibrium determination of such rental income.

Consider the function:

$$Z(t, D_A, D_B) = \frac{K}{n_R}\left(Q_K\left(K, f(\alpha(t + D_A, n_{RA}))n_{PA} + f(\alpha(t + D_B, n_{RB}))n_{PB}\right)\right). \tag{3.2}$$

We have, from (3.2), (since $Q_{KL}, \alpha_I, f' > 0$):

$$Z_t = \frac{K}{n_R}(Q_{KL})\left(f'(\alpha(t + D_A, n_{RA}))\alpha_I(t + D_A, n_{RA})n_{PA} + f'(\alpha(t + D_B, n_{RB}))\alpha_I(t + D_B, n_{RB})n_{PB}\right) > 0, \tag{3.3}$$

and, for every $k \in \{A, B\}$,

$$Z_{D_k} = \frac{K}{n_R}(Q_{KL})\left(f'(\alpha(t + D_k, n_{Rk}))\alpha_I(t + D_k, n_{Rk})n_{Pk}\right) > 0. \tag{3.4}$$

Since, by assumption, $Q_{KLL} \leq 0, f'' < 0, \alpha_{II} \leq 0$,, it follows from (3.3) that:

$$Z_{tt}, Z_{tD_k} < 0. \tag{3.5}$$

Note now that $Z(0,0,0) = \frac{K}{n_R}(Q_K(K, n_P)) > 0$. Let $\hat{t} = \left[\frac{K}{n_R}(Q_K(K, n_P))\right]$.

Then, since $D_A, D_B \geq 0$, (3.3)-(3.4) together imply:

$$Z(\hat{t}, D_A, D_B) > \hat{t}. \tag{3.6}$$

We assume the following.

A6. $Z_t(\hat{t}, 0, 0) < 1$.

In light of (3.3) and (3.6), then, the next result is immediate.

Lemma 3.6. *Let* $Z(t, D_A, D_B)$ *be defined by (3.2), and let* $\hat{t} = \left[\frac{K}{n_R}(Q_K(K, n_P))\right]$. *Given A6 and* (D_A, D_B),

(i) there exists a unique $t^* \in \Re_+$ *such that: (a)* $Z(t^*, D_A, D_B) = t^*$, *(b) for all* $t \in [0, t^*)$, $Z(t, D_A, D_B) > t$, *and (c) for all* $t > t^*$, $Z(t, D_A, D_B) < t$;

(ii) furthermore, $t^* > \hat{t}$, *and, for every* $k \in \{A, B\}$, $t^*_{D_k} > 0$.

Lemma 3.6 is illustrated in Figure 1 below. The arrow shows how the Z schedule moves up with an increase in D_k.

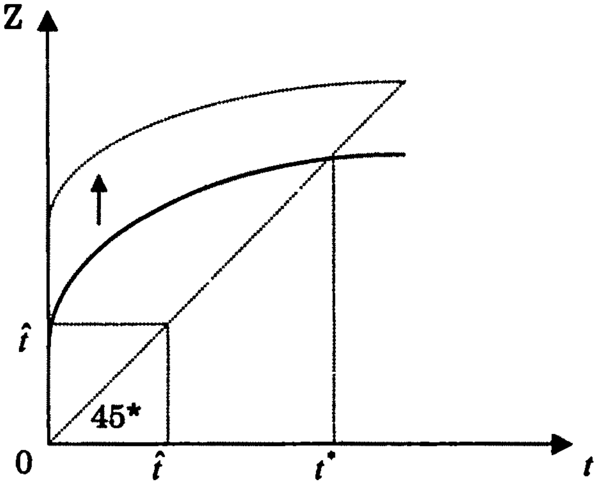

Figure 1

Our last assumption is required to ensure that, given the amounts of foreign flows received by R agents, D_A, D_B, the corresponding value of $t^*(D_A, D_B)$ can be achieved as rental income, given the capital stock, K, in the economy.

A7. $t^*(D_A, D_B) < \frac{K}{n_R} Q_K(K, \delta n_P)$.

We are now ready to address the issue of general equilibrium in our economy. Suppose A1-A7 are all satisfied. Since $Q_{K\lambda} > 0$, in light of Lemma 3.6, A6-A7 imply:

$$\lfloor n_P < f(\alpha(t^* + D_A, n_{RA}))n_{PA} + f(\alpha(t^* + D_B, n_{RB}))n_{PB} < \delta n_P \rfloor.$$

Then, $\lfloor w^* = d(f(\alpha(t^* + D_A, n_{RA}))n_{PA} + f(\alpha(t^* + D_B, n_{RB}))n_{PB})\rfloor$ is well-defined. Consider w^*. Given A1-A5, by Lemmas 3.1, 3.3 and 3.4, if the rental income of every R individual is t^*, then labor supply from community k, in the Nash equilibrium corresponding to $\langle t^* + D_k, w^* \rangle$, is $f(\alpha(t^* + D_k, n_{Rk}))n_{Pk}$; thus, total labor supply is $(f(\alpha(t^* + D_A, n_{RA}))n_{PA} + f(\alpha(t^* + D_B, n_{RB}))n_{PB})$. Hence, w^* constitutes an equilibrium wage rate in the labor market if the rental income of every R individual is t^*. Given the employment at this equilibrium, $(f(\alpha(t^* + D_A, n_{RA}))n_{PA} + f(\alpha(t^* + D_B, n_{RB}))n_{PB})$, capital market equilibrium indeed yields a rental income of t^* per R individual in community k. Hence, a general equilibrium exists in our economy. Lemma 3.6 (i) also implies that this general equilibrium must be unique (and stable).

We now combine Lemma 3.1, Lemma 3.3, Lemma 3.4 and Lemma 3.6 to present the following proposition.

Proposition 3.7. *Given A1-A7, there exists a unique* $w^* \in [\underline{w}, \overline{w}]$ *such that labor supply generated by community A in the Nash equilibrium corresponding to* $\langle t^* + D_A, w^* \rangle$, *and that generated by community B in the Nash equilibrium corresponding to* $\langle t^* + D_B, w^* \rangle$, *together add up to exactly* $d^{-1}(w^*)$, *where* $t^* = \frac{K}{n_R} Q_K(K, d^{-1}(w^*))$. *Furthermore, these community-specific Nash equilibria are such that:*

(i) $d^{-1}(w^*) = \lfloor f(y_A^*)n_{PA} + f(y_B^*)n_{PB} \rfloor$, *where* y_A^*, y_B^* *are the amounts of the public goods specific to communities A and B, respectively, generated in the Nash equilibria of the respective communities;*

(ii) no R agent enters the labor market,

(iii) every P agent belonging to community k supplies $f(y_k^*)$ *efficiency units of labor,*

(iv) no P agent contributes to her community's public good,

(v) for all $k \in \{A,B\}$, $y_k^* = \alpha(t^* + D_k, n_{Rk})$,

and

(vi) w^* *decreases in* D_A, D_B.

We can now proceed to address the issue of comparative statics.

4. Comparative statics

Consider an exogenous increase in D_k, i.e., the amount of foreign flow to capitalists, provided in the form of the single commodity in our economy. Since there is no investment in our model, capitalists will consume all such additional income, the increment in demand being supplied from abroad. Consequently, there will be no change in the demand for labor schedule. It follows that, in the standard framework, where workers' labor supply does not depend on capitalists' consumption, workers' earnings will remain invariant. Thus, additional income for capitalists will lead to a Pareto improvement. What happens in our framework?

Proposition 3.7 (vi) implies that, if foreign flow incomes of capitalists increase even in just one community, then the equilibrium wage rate per efficiency unit of labor must fall. Since the labor demand schedule remains unchanged, this implies that aggregate labor supply, in efficiency units, must rise. Consequently, rental incomes of *all* capitalists must rise as well. Thus, a prior increase in the incomes of capitalists even in one community sets in motion community and market processes that generate additional gains for every capitalist, regardless of community affiliation. Capitalists gain as a class, by virtue of their ownership of capital assets. What happens to the wage income, and welfare, of an individual worker?

Let aggregate employment in the initial equilibrium be λ^*, total labor supplied by community k be λ_k^*, and the wage rate w^*. Then, by Proposition 3.7((ii) and (iii)), the equilibrium wage income of a P individual, belonging to community k, is given by:

$$M_{Pk}^* = w^* f(y_k^*) = \frac{\lambda_k^* w^*}{n_{Pk}}. \qquad (4.1)$$

Given any community k, let –k denote the other community. Then, (4.1) yields:

$$\text{for every } j \in \{A,B\},\ \frac{\partial M^*_{Pk}}{\partial D_j} = \frac{w^*}{n_{Pk}}\left[\frac{\partial \lambda^*_k}{\partial D_j}\right]\left[1 - \frac{\theta^*_k}{\eta^* \Delta^*_k}\right],$$

$$\text{where: } \eta^* = -\frac{w^*}{\lambda^* d'(\lambda^*)}, \theta^*_k = \left[\frac{\lambda^*_k}{\lambda^*}\right] \text{and } \Delta^*_k = \left[\frac{\frac{\partial \lambda^*_k}{\partial D_j}}{\frac{\partial \lambda^*}{\partial D_j}}\right]. \tag{4.2}$$

Note now that Proposition 3.7((ii), (iii) and (v)) together yield $\left[\frac{\partial \lambda^*_k}{\partial D_j}\right]$, $\left[\frac{\partial \lambda^*_{-k}}{\partial D_j}\right]$>0. Using (4.2), we thus have the following.

Lemma 4.1. *For every* $j \in \{A,B\}$, $\frac{\partial M^*_{Pk}}{\partial D_j} < 0$ *if and only if* $\left[\eta^* \Delta^*_k < \theta^*_k\right]$.

Since $\Delta^*_k \in (0,1)$, by Lemma 4.1, a small increase in foreign income of every capitalist in at least one community must necessarily reduce the income of every worker in community k if the elasticity of labor demand at the initial equilibrium is less than the share of community k in total labor supply. If the elasticity of labor supply at the initial equilibrium is less than the share of either community in total labor supply, i.e., if $\eta^* < \min\{\theta^*_A, \theta^*_B\}$, then a small increase in foreign income of every capitalist in at least one community must necessarily reduce the incomes of *all* workers in society. More generally, if labor demand is inelastic, then such an increase must *necessarily* reduce the earnings of workers belonging to *at least one* community. This is possible (though not necessary) even when labor demand is elastic. Intuitively, this will happen if the productivity of workers in one community goes up much less than the productivity of their counterparts in the other community (Δ^*_k is low). This in turn is likely when (a) capitalists of the other community, -k, receive the wealth increase, and/or (b) community k already has a large stock of the public good (f'<0).[18]

Since P individuals are non-contributory to the public good, it must be the case that, if any given increase in supply of the public good is matched by a reduction in wage income of the same amount, then the welfare of these individuals must fall.

---

[18] In this case, however, workers in at least one community must see their incomes rise.

Thus, from (4.2), we have a sufficient condition for a reduction in welfare of P agents belonging to community k:

$$\left[1+w^{*}f'(y_k^{*})\left[1-\frac{\theta_k^{*}}{\eta^{*}\Delta_k^{*}}\right]\right]<0. \tag{4.3}$$

It is evident that (4.3) will indeed hold if either the elasticity of labor demand, or the share of community k in additional labor supply, is sufficiently low.

We summarize our results in the form of the following Proposition.

Proposition 4.2. *Suppose A1-A7 are satisfied. Let the share of community k in total labor supply in efficiency units at the initial labor market equilibrium be* θ_k^{*}*, and let the elasticity of labor demand at that equilibrium be* η^{*}*. Consider a small increase in foreign income of every R agent in at least one community, and let* Δ_k^{*} *denote the share of community k in the additional labor supply generated in the equilibrium subsequent to this increase. There exists* $\hat{\eta}\in(0,\theta_k^{*})$ *such that, if* $\eta^{*}\Delta_k^{*}<\hat{\eta}$*, then the welfare of every P agent in community k must be lower in the new equilibrium.*

Lastly, how does an increase in capitalists' incomes affect inequality between workers belonging to different communities? Consider the special case where $n_{RA}=n_{RB}=\tilde{n}$. From (4.1), we have:

$$\frac{\partial M_{PA}^{*}}{\partial D_B}=f(y_A^{*})\frac{\partial w^{*}}{\partial D_B}+w^{*}f'(y_A^{*})\alpha_I(t^{*}+D_A,\tilde{n})t_{D_B}^{*}. \tag{4.4}$$

$$\frac{\partial M_{PB}^{*}}{\partial D_B}=f(y_B^{*})\frac{\partial w^{*}}{\partial D_B}+w^{*}f'(y_B^{*})\alpha_I(t^{*}+D_B,\tilde{n})(1+t_{D_B}^{*}). \tag{4.5}$$

$$\frac{\partial M_{PA}^{*}}{\partial D_B}+\frac{\partial M_{PA}^{*}}{\partial D_A}=f(y_A^{*})\left(\frac{\partial w^{*}}{\partial D_B}+\frac{\partial w^{*}}{\partial D_A}\right)+w^{*}f'(y_A^{*})\alpha_I(t^{*}+D_A,\tilde{n})(t_{D_B}^{*}+t_{D_A}^{*}+1). \tag{4.6}$$

First note that, if, in the initial equilibrium, A workers earn more than B workers, then it must be that $D_A>D_B$. Since, by assumption, $f''<0,\alpha_{II}\leq 0$, it follows that, if $M_{PA}^{*}\geq M_{PB}^{*}$, then an increase in foreign incomes of R agents in B alone increases the incomes of P agents in B more than their counterparts in A, and thereby reduces income inequality among workers. Paradoxically, if initially P agents in A earn significantly less than their counterparts in B, then the exogenous increase in the incomes of R agents in B can increase the incomes of P agents in A more than those in B, and thereby, once again, reduce earnings inequality across communities among the poor. On the other hand, in intermediate situations where initially P agents in A earn somewhat less than those in B, such an increase will aggravate inequality among the poor. If all R agents receive identical increases in foreign income, then those P agents who were initially better off must gain less, or lose more; thus, there will be an unambiguous reduction in income inequality amongst workers across communities.[19]

5. Conclusion

In this paper, we have shown how a prior increase in the incomes of capitalists, or, more generally, rich individuals, sets in motion community and competitive market processes which subsequently (a) further increase capitalists' incomes, (b) can reduce workers' earnings as well as welfare, and (c) systematically influence earnings differentials between workers belonging to different communities. In doing so, we have presented an integrated, general equilibrium, analysis of how the presence of vertical ties of 'community' between sections of workers and sections of capitalists can critically affect distribution of income between capitalists as a class and workers as a class, as well as between workers belonging to different communities.

We have modeled 'community' in terms of a group-specific public good, produced according to the standard, additive, technology. This however is purely for convenience of exposition. More general specifications of public good technology, such as the CES functional form (Cornes 1993) and 'impure' public goods, where agents derive utility not only from the total amount of the public good, but from the size of their own contributions as well, are compatible with our analysis. In assuming a one-period framework, we have abstracted from issues surrounding investment. In a more general dynamic setting, one would expect an increase in incomes of capitalists to be partially reflected in increased investment, and thereby, additional labor demand.

Our central point, however, is that when community intersects with class, the standard analysis of distribution is modified significantly, and in interesting ways. We look forward to further explorations of this intersection in the economics literature.

---

[19] These claims will hold in the more general case, with different numbers of R agents in the two communities, as well so long as $M^*_{Pk} > M^*_{P(-k)}$ implies

$$\alpha'\left(t^* + D_k, n_{Rk}\right) \leq \alpha'\left(t^* + D_{-k}, n_{R(-k)}\right).$$

Appendix

To prove Lemma 3.3, we first note the following.

Lemma N1. *Given A1, any* $k \in \{A, B\}$, *any* $I_{Rk} \geq 0$ *and any* $w \in [\underline{w}, \overline{w}]$, *every R agent must contribute a positive amount to the public good of her own community in any Nash equilibrium corresponding to* $\langle I_{Rk}, w \rangle$.

Proof of Lemma N1.

Suppose a Nash equilibrium exists corresponding to $\langle I_{Rk}, w \rangle$. It is easy to check that at least one class of agents must always contribute to the public good of the community in a Nash equilibrium. Hence, to establish the claim above, we only need to show that, if P agents are contributory in the Nash equilibrium, then R agents must be contributory in the Nash equilibrium as well. Suppose not. Let x_{Rk}, x_{Pk} be the private consumptions of R and P individuals,

respectively, in community k. Then, since $\frac{g_x}{g_y}$ is decreasing in x_i by A1(ii) (Remark 2.1), it must be the case that, in the Nash equilibrium, $x_{Rk} < x_{Pk}$. In an equilibrium, however, for every individual, $wfu_x \leq 1$. If $x_{Rk} < x_{Pk}$ in equilibrium, then, given $u_{xx} < 0$ (A1(ii)), it must be the case that $wf(y^k)u_x(x_{Pk}, y^k) < 1$ in equilibrium. But this implies that P agents provide 0 market labor. As $I_{Pk} = 0$, this in turn implies that P agents cannot contribute to the community's public good, a contradiction which establishes our claim. ◊

Proof of Lemma 3.3.

Let the amount of the public good of community k, generated in the Nash equilibrium be y^*, the amount of private good consumed by an R individual belonging to k, *i*, be x_i^* and the amount of labor supplied by this individual be L_i^*. Since A1 holds by assumption, by Lemma N1, R individuals in k must contribute to the public good of k. Then the Nash equilibrium is given as the solution to the problem (P1) subject to (2.2) alone. Given , is defined as the solution to the following pair of equations:

$$\frac{g_x(x_i^*, y^*)}{g_y(x_i^*, y^*)} = \frac{1}{1 - wf'(y^*)L_i^*}, \tag{N1}$$

and

$$x_i^* + y^* = I_{Rk} + wf(y^*)L_i^* + y_{-i}^*. \tag{N2}$$

Let $p^* = (1 - wf'(y^*)L_i^*)$. The budget constraint (N2) above can then be rewritten as:

$$x_i^* + p^* y^* = I_{Rk} + y_{-i}^* + wL_i^* y^* \left(\frac{f(y^*)}{y^*} - f'(y^*) \right) = r^*. \tag{N3}$$

It follows that $\langle x_i^*, y^* \rangle$ constitutes the solution to:

$\underset{x_i, y^k}{Max}\ g(x_i, y^k)$ subject to: $x_i + p^* y^k = r^*$,

which implies:

$$g_x(x_i^*, y^*) = V_r(r^*, p^*). \tag{N4}$$

Note now that, since $f'' < 0$, (N3) implies $r^* > I_{Rk}$. Then, since $p^* \in [\underline{p}, 1]$, using A3 and (N4), we have:

$$g_x(x_i^*, y^*) < \frac{1}{\overline{w}\delta}.$$

Since $wf(y^*) < \overline{w}\delta$, we thus get:

$$g_x(x_i^*, y^*) < \frac{1}{wf(y^*)}. \tag{N5}$$

(N5) however implies $L_i^* = 0$. ◊

G165642

References

Akerlof, George A. and Rachel E. Kranton (2000): "Economics and Identity," *Quarterly Journal of Economics*, 115(3): pp. 715-53.

Alesina, Alberto, Reza Baqir and William Easterly (1999): "Public Goods and Ethnic Divisions", *Quarterly Journal of Economics*, 114(4): pp. 1243-1284.

Alesina, Alberto and Dani Rodrik (1994): "Distributive Politics and Economic Growth", *Quarterly Journal of Economics*, 109(2): pp. 465-490.

Anand, Sudhir (1983): *Inequality and Poverty in Malaysia: Measurement and Decomposition*, Oxford: Oxford University Press.

Bergstrom, T.C., Larry Blume and Hal Varian (1986): "On the Private Provision of Public Goods," *Journal of Public Economics*, 29, pp. 25-49.

Cornes, Richard (1993): "Dyke Maintenance and Other Stories: Some Neglected Types of Public Goods," *Quarterly Journal of Economics*, 108(1): pp. 259-271.

Dasgupta, Indraneel and Ravi Kanbur (2001): "Class, Community, Inequality," WP 2001-24, Dept. of Applied Economics and Management, Cornell University.

Fafchamps, Marcel and Bart Minten (2002): "Returns to Social Network Capital Among Traders." *Oxford Economic Papers* 54: 173-206.

Kanbur, Ravi (2000): "Income Distribution and Development," in A.B. Atkinson and F. Bourguignon (eds.) *Handbook of Income Distribution*, Amsterdam: Elsevier-North Holland.

Kanbur, Ravi and Nora Lustig (2000): "Why is Inequality Back on the Agenda?" *Proceedings of the Annual World Bank Conference on Development Economics*, Washington, DC: World Bank.

Lancaster, Kelvin (1973): "The Dynamic Inefficiency of Capitalism", *Journal of Political Economy*, 81(5): pp. 1092-1109.

Ligon, Ethan, Jonathan Thomas and Tim Worrall (2002): "Informal Insurance Arrangements in Village Economies," *Review of Economic Studies*, forthcoming.

Loury, Glenn (2002): *The Anatomy of Racial Inequality*, Cambridge, MA: Harvard University Press.

Proudhon, Pierre-Joseph (1994): *What is Property?* Cambridge Texts in the History of Political Thought, Cambridge: Cambridge University Press.

Przeworski, Adam and Michael Wallerstein (1982): "The Structure of Class Conflict in Democratic Capitalist Societies", *American Political Science Review*, 76(2): pp. 215-238.

Roemer, J.E. (1998): "Why the Poor do not Expropriate the Rich: An Old Argument in New Garb", *Journal of Public Economics*, 70(3): pp. 399-24.

Somanathan, E. (2001): "Can Growth Ease Class Conflict?" forthcoming in *Economics and Politics*.

Wetherell, Margaret (1996): "Group Conflict and the Social Psychology of Racism," in Margaret Wetherell (editor), *Group Conflict and the Social Psychology of Racism*, Thousand Oaks, CA: SAGE Publications.

Arthaniti 5 (1-2), 24-33 / 2006

On the Incentives for Cooperative and Non-Cooperative R&D in Duopoly

Tarun Kabiraj*

Abstract

This paper studies incentives for cooperative research vis-à-vis non-cooperative research when there are two conceivable product innovations. We show that, contrary to the literature, there are situations where non-cooperative R&D generates a larger payoff even when the success probability in R&D is high. We consider different forms of cooperation in this context. If research cooperation has access to all these forms, cooperative R&D will dominate at least for high probability of success. We have also studied the effect of patent protection on the choice of R&D institution. It is shown that patent protection reduces the scope of cooperation in R&D.

Keywords : Cooperative R&D, non-cooperative R&D, patent protection, imitation.

JEL Classification : D43, L13, O31.

1. Introduction

Research and development (R&D) is considered to be the single-most important activity for industrial growth and economic development of a country. But this is often constrained by various factors. Such an activity involves huge expenditures in the form of setting research lab, installing scientific instruments and employing scientific personnel. Moreover, the research outcomes are uncertain, and sometimes the failure rate is quite high. Even when an innovation occurs successfully, the innovator may fail to appropriate the requisite amount of profits because of spillovers, leaking out of knowledge and free-riding problems. To the extent patent protection prevents imitation and leaking out of knowledge, it gives incentives to the investors for doing R&D, but obviously, it cannot solve the whole incentive problem. So private investors may not get sufficient incentives for doing R&D. Even many socially valuable innovations may not be taken up ever by the private investors. As a result, the actual R&D activity may be far less than the desired level. In such a situation, cooperative research is thought to be a way out and hence cooperative research is often encouraged.

There is already a wide literature that deals with the choice of R&D institution under different scenarios.[1] This literature has focused on the interaction of firms competing

---

* Tarun Kabiraj, Economic Research Unit, Indian Statistical Institute, 203, B. T. Road, Kolkata 700108, India. E-mail: tarunkabiraj@hotmail.com; Fax: 91-33-25778893.

[1] A subset of the literature comprises Katz (1986), D'Aspremont and Jacquemin (1988), Marjit (1991), Combs (1992), Kamien, Muller and Zang (1992), Suzumura (1992), Choi (1992, 1993), Kabiraj and Mukherjee (2000), and Silipo and Weiss (2005).

for a single innovation. In contrast, in the present paper we consider R&D interaction of two firms when there are two conceivable product innovations. While R&D can be cooperative or non-cooperative, production is assumed to be always non-cooperative. The firms decide their investment projects and accordingly they invest. But the R&D outcomes are uncertain, therefore, the success or failure in an R&D activity occurs with a positive probability. We discuss the problem under different scenarios, e.g., patent protection may or may not be available, imitation of rival's production knowledge is possible or not, etc. We assume that each firm has its own research lab, and each lab has the capacity to take up at most one research project. So we consider the situations when the research cooperation decides to do R&D in a single lab (in which case the partners share both knowledge and R&D investment) or in both labs (in which case the cooperation decides to take up the same innovation project or different innovation projects).

The present paper has the framework of Marjit (1991) and Combs (1992), but it differs from theirs in various respects. First, as we have already mentioned, they have considered the possibility of a single innovation only. In contrast, we have two possible innovations. Second, the underlying implicit assumption in their models is that either the firms cannot imitate rivals' innovations at all (so that patenting issue is irrelevant), or all the firms successful in research get patent protection. Neither of these assumptions seems realistic. So we have discussed the problem under different assumptions. When imitation is not at all possible (equivalently, imitation is too costly), patent protection issue is not important, but when imitation is costless, whether patent protection is available and effective is an important consideration. In the present paper we have studied the effect of patent protection on the choice of R&D institution. We have shown that the availability of patent protection reduces the scope of R&D cooperation. To the extent patent protection gives incentives for R&D investment, it partly takes away the incentives for doing cooperative R&D.

We have also contrasted our results with Marjit and Combs. Introducing the possibility of two innovations in an otherwise Marjit model drastically alters the result. We have shown that, contrary to the literature, non-cooperative R&D can occur even when the probability of success in R&D is relatively large. Our result follows in a situation when the research cooperation decides to do R&D in a single lab (and so takes up a single innovation project). We have therefore discussed the other forms of cooperation. We have shown that doing research on two different innovation projects in two different labs dominates all other forms of cooperation, and under this situation cooperative research also dominates non-cooperative research at least for high probability of success.[2]

---

[2] In Combs (1992) there are m research projects for a single product innovation, and exactly one of these is successful. And the research joint venture is supposed to select two projects without replacement.

The setup of the paper is as follows. The second section provides the model and results of the paper. The third section is a conclusion.

2. Model

Consider two firms, 1 and 2, interacting in R&D and production. While R&D can be cooperative or non-cooperative, production is assumed to be always non-cooperative. In the product markets, firms compete in quantities.

There are only two conceivable products, X and Y, which could be innovated by well-directed R&D activities. Each such R&D project involves an investment of resources of an amount $R > 0$. However, the results of an R&D activity are uncertain; each such project yields a successful outcome with a probability, p, $0 < p < 1$. Both p and R are unrelated and exogenously specified. To simplify the analysis, we assume that two products are independent. We further assume that the market demands for these products are identical and same.

We have the following game structure. First, the firms decide whether R&D activity will be conducted cooperatively or non-cooperatively. Given the decision, they decide which R&D project(s) to be selected. They accordingly invest in R&D. We assume that each firm has already a research lab, and each lab has the capacity of conducting at most one R&D project. Given the outcomes of the R&D stage, the firms then compete non-cooperatively in the product market. Each product, if it gets innovated, has the same cost of production function. We assume that both the firms are risk-neutral. In the following two sub-sections we shall discuss the outcomes of the non-cooperative and cooperative R&D activities under different assumptions. Then in the third sub-section we discuss the choice between cooperative and non-cooperative research.

2.1 Non-cooperative R&D

Under non-cooperative R&D, each firm first independently decides which project, X or Y, will be taken up. Given the choice of the project, a firm believes that the other firm will have the same project choice with probability ½ and the different project choice with probability ½. Thus the probability that both the firms have targeted the same product innovation is ½ and that they have different R&D projects is ½. Each such project will yield a successful outcome with probability p. However, the expected payoff of a firm will depend, as we see below, on the assumption of whether imitation of an innovation is possible and/or whether patent protection for such an innovation is available. We consider the following scenarios.

Case 1: Patent protection available

If under this situation the firms have different R&D projects, then any firm which comes up with a success (this occurs with probability p) gets patent protection for its innovation and hence will enjoy monopoly in the product market for this product. If both the firms have invested on the same R&D project and only one of them is successful (this occurs with probability $p(1-p)$), the successful firm gets the patent. But if both firms are successful to innovate (which occurs with probability p^2), each firm is assumed to get protection with probability one-half (1/2). Denoting the monopoly payoff by π^m, the expected payoff of a firm under this case will be given by:

$$E(N1) = \frac{1}{2}[p\pi^m + (1-p).0] + \frac{1}{2}[\frac{p^2}{2}\pi^m + p(1-p).\pi^m] - R$$

The first term of the expression on the right hand side denotes the expected net revenue when they have different project choices, and the middle term is the expected profit when they have the same R&D project. The expression can be simplified to get:

$$E(N1) = \frac{1}{4}p(4-p)\pi^m - R \tag{1}$$

Case 2: Neither patent protection nor imitation possible

Assume that patent protection is not available (or equivalently, patent protection is ineffective), but at the same time imitation of other's innovation is not possible.[3] Therefore, when the firms come up with the same innovation, the market structure is symmetric duopoly, whereas when only one firm has successful innovation, it has monopoly. If they have different innovations, each has monopoly in the respective market. Under this situation, the expected payoff of a firm is:

$$E(N2) = \frac{1}{2}p\pi^m + \frac{1}{2}[p^2\pi^d + p(1-p).\pi^m] - R$$

$$\text{or, } E(N2) = \frac{1}{2}p[(2-p).\pi^m + p\pi^d] - R \tag{2}$$

where π^d is the symmetric duopoly profit in a market.

Case 3: Costless imitation but no patent protection

Consider the scenario where patent protection is not available, but at the same time each firm can perfectly imitate the rival's innovation at zero cost. Equivalently, this is the case of perfect spillover of knowledge. Thus, contrast to the previous case, the market structure will be duopoly for each product if at least one firm gets the innovation. Hence the expected payoff of each firm is:

---

[3] These were the underlying assumptions in Marjit (1991) and Combs (1992). However, they do not have the possibility of two innovations.

$$E(N3)=\frac{1}{2}[p^2.2\pi^d+2p(1-p).\pi^d]+\frac{1}{2}[p^2\pi^d+2p(1-p).\pi^d]-R$$

or, $$E(N3)=\frac{1}{2}p(4-p).\pi^d-R \qquad (3)$$

Throughout our analysis we assume $\pi^m > 2\pi^d$, that is, monopoly payoff is greater than the industry duopoly profit.[4] Now, comparing (1), (2) and (3), we have:

$$E(N1)>E(N2)>E(N3) \qquad (4)$$

This simply tells that the availability of patent protection increases R&D incentives, and the possibility of imitation and spillovers reduces incentives for R&D.

2.2 Cooperative R&D

In this subsection we consider that the firms do R&D cooperatively. Given our structure, the research cooperation then decides whether to do research in both labs and share the results between the firms, or whether to do in a single lab and share information as well as R&D cost. Obviously, when the research cooperation decides to do research in a single lab, the constituent members invest only on a single project (either X or Y). But if they decide to do research in both labs, then either both research labs will take up the same project, or one lab will target X innovation and the other lab Y innovation. In any case, if any lab successfully innovates the product, both firms will share the knowledge. Note that, no matter whether there is imitation or patent protection, the market structure for each product, if successfully innovated, will be symmetric duopoly. Therefore, the expected payoffs under different scenarios of R&D cooperation will be as follows.

Case (i): When the firms decide to do R&D in a single lab and share the R&D cost (and hence take up a single project), the expected payoff of each firm is:

$$E(C1)=p\pi^d-\frac{R}{2} \qquad (5)$$

This is in fact the payoff as could be derived from Marjit (1991) if it is reworked for product innovation.

Case (ii): When the firms take up the same project in both labs, the corresponding expected payoff expression is:

---

[4] This is obviously true for the case of linear cost function. For the linear demand function, $P=a-b(q_1+q_2)$ and quadratic cost function $C=\alpha q_i^2$, $b,\alpha>0$, we have $\pi^m > 2\pi^d$ if and only if $b>2(1+\sqrt{2})\alpha$.

$$E(C2) = [p^2 + 2p(1-p)]\pi^d - R$$

$$\text{or, } E(C2) = p(2-p)\pi^d - R \tag{6}$$

Case (iii): Finally, if they consider two different research projects in two labs, the expected payoff will be given by:

$$E(C3) = [p^2.2\pi^d + 2p(1-p)\pi^d] - R$$

$$\text{or, } E(C3) = 2p\pi^d - R \tag{7}$$

Combs (1992) came up with the same payoff expression as (7). While Combs considered single product innovation, but each firm was allowed to choose different projects targeting the same product innovation.

Generally, all these different forms of cooperation are not available to the research cooperation because of coordination and moral hazard problem.[5] Since we have not addressed these issues in the present paper, we assume that all these forms of cooperation are feasible and available. Hence the research cooperation will choose the optimal form. Comparing (5), (6) and (7), we have

$$E(C3) > max\{E(C1), E(C2)\} \tag{8}$$

Thus doing R&D on two different projects in two labs is optimal in the present model. However,

$$E(C1) \gtreqless E(C2) \Leftrightarrow (R/2) \gtreqless p(1-p)\pi^d \tag{9}$$

that is, for relatively small and large probability of success, $C1$ will dominate $C2$ and for middle level of probabilities, $C2$ will dominate $C1$.

2.3 Cooperative vs. Non-cooperative Research

We are now in a position to decide the choice between the two forms of R&D institution. Consider first the situation when patent protection is available. Hence comparing (1) and (7) we shall get:

$$E(C3) > E(N1) \Leftrightarrow p > p^* = \frac{4[\pi^m - 2\pi^d]}{\pi^m} \tag{10}$$

$$\text{And} \quad p^* < 1 \text{ if and only if } \frac{\pi^d}{\pi_m} > \frac{3}{8} \tag{11}$$

CALCUTTA UNIVERSITY CENTRAL LIBRARY

---

[5] See Choi (1992) and Kamien et al. (1992) in this context.

This gives our first proposition.

Proposition 1: *Assume that patent protection is available and* $\frac{\pi^d}{\pi_m} > \frac{3}{8}$. *Then* $\exists p^* < 1$ *such that* $\forall p \in (p^*, 1)$, *cooperative R&D will give a larger incentive for innovation; non-cooperative research should be preferred if* $p \leq p^*$.

Obviously, the underlying condition of the result holds for a wide class of demand and cost functions. For instance, for linear demand and cost functions, we have, $\frac{\pi^d}{\pi_m} > \frac{4}{9}$.

Now consider the situation when neither patent protection is available, nor imitation is possible. Then cooperative R&D will generate a larger payoff if and only if

$$E(C3) > E(N2) \Leftrightarrow p > \tilde{p} = \frac{2[\pi^m - 2\pi^d]}{\pi^m - \pi^d} \tag{12}$$

Note that

$$\tilde{p} < 1 \Leftrightarrow \frac{\pi^d}{\pi_m} > \frac{1}{3} \text{ and } \tilde{p} < p^* \tag{13}$$

Here cooperative R&D dominates non-cooperative R&D for $p \in (\tilde{p}, 1)$. Therefore, when patent protection is not available, the possibility of cooperation goes up.

Finally consider the choice of the research institution when perfect imitation is possible, but patent protection is not available.[6] Then comparing (3) and (7), it is easy to see that

$$E(C3) > E(N3) \forall p > 0 \tag{14}$$

That is, under this situation cooperative research always yields a larger payoff, given that R&D outcomes are uncertain.

Summarizing the above discussion we see that *if the probability of success in research is relatively large, and if the research cooperation chooses the optimal form of cooperation, cooperative R&D will strictly be preferred to non-cooperative R&D.*

---

[6] Equivalently, this is the scenario where there are spillovers of knowledge and patent protection is ineffectively enforced.

The effect of patent protection on the choice of R&D institution is captured in the following proposition.

Proposition 2: *Patent protection reduces the scope of cooperation in R&D.*

Proof: To prove this, consider situations when cooperative R&D would occur in the absence of patent protection, but non-cooperative R&D occurs under patent protection. Thus consider any $p \in (\tilde{p}, p^*)$. For such a p, we have, $E(N1) > E(C3) > E(N2)$. Hence the result. ♦

As an implication of the above result it follows that if patent protection is ineffective to prevent imitation and spillover of knowledge, cooperative R&D can be an effective solution to the incentive problem in such a situation. For instance, think of the scenario when $E(C3) > 0 > E(N3)$. In this case, there will be no R&D investment under non-cooperative situation, whereas it will be feasible under cooperation.

Before we conclude this section, let us see how Marjit (1991) result will be altered with the introduction of two potential innovations. Recall that in Marjit, research cooperation occurs in a single lab and the partners share both knowledge and R&D investment. Further, both patent protection and imitation are absent. Hence we compare $E(C1)$ and $E(N2)$. We find that

$$E(C1) >,< E(N2) \Leftrightarrow \frac{R}{2(\pi^m - \pi^d)} >,< (p - p^2/2) \tag{15}$$

The expression in the extreme right hand side, $\phi(p) = p - p^2/2$, is increasing and concave in p, with $\phi(0) = 0$ and $\phi(1) > 0$. Hence, if R is not too large, non-cooperative R&D will occur for relatively large p. If in this otherwise Marjit model with two potential innovations, we now introduce patent protection, the incentives for non-cooperative R&D will further go up. Thus, in contrast to the existing literature, we have derived the following result.

Proposition 3: *There are situations where non-cooperative R&D will occur if and only if the probability of success in research is large.*

3. Conclusion

This paper deals with the choice of R&D institution in a duopoly. The existing literature has addressed this problem in the context of a single innovation. In contrast, we consider two potential innovations. However, success in R&D is probabilistic. The payoff of a firm under non-cooperative R&D depends on whether imitation of an innovation is at all possible and whether patent protection is available. And the payoff under cooperative R&D depends on the form of cooperation, that is, whether R&D occurs in a single lab or in both labs and whether the firms invest on the same R&D project or on different

R&D projects. If the research cooperation operates on a single lab sharing both R&D cost and knowledge, then we have shown that, contrary to the literature, non-cooperative R&D generates a larger payoff if the probability of success is large. On the other hand, if the research cooperation can choose any other forms of cooperation, cooperative R&D dominates non-cooperative R&D at least for large probability of success. Thus the choice of an R&D institution depends on the available forms of cooperation and on the extent of patent protection. We have also studied the effect of patent protection on the incentive for cooperative research. It is shown that the availability of an effective patent system reduces the scope of cooperative R&D. So in a situation when patent protection is ineffective, cooperative research will necessarily enhance incentives for R&D investment, because then it avoids the negative effect of imitation and spillovers.

References

Choi, J.P. (1992), "Cooperative R&D with moral hazard", *Economics Letters* 39, 485-491.

Choi, J.P. (1993), "Cooperative R&D with product market competition", *International Journal of Industrial Organization* 11, 553-571.

Combs, K.L. (1992), "Cost sharing vs. multiple research projects in cooperative R&D", *Economics Letters* 39, 353-357.

D'Aspremont, C. and A. Jacquemin (1988), "Cooperative and non-cooperative R&D in duopoly with spillovers", *American Economic Review* 5, 1133-37.

Kabiraj, T. and A. Mukherjee (2000), "Cooperation in R&D and production — A three-firm analysis", *Journal of Economics* 71, 281-304.

Kamien, I.M., E. Muller and I. Zang (1992), "Research joint ventures and R&D cartels", *American Economic Review* 82, 1293-1306.

Katz, M.L. (1986), "An analysis of cooperative research and development", *Rand Journal of Economics* 4, 527-543.

Marjit, S. (1991), "Incentives for cooperative and non-cooperative R and D in duopoly", *Economics Letters* 37, 187-191.

Silipo, D.B. and A. Weiss (2005), "Cooperation and competition in a duopoly R&D market", *Research in Economics* (forthcoming)

Suzumura, K. (1992), "Cooperative and non-cooperative R&D in an oligopoly with spillovers", *American Economic Review* 82, 1307-20.

Arthaniti 5 (1-2), 34-63 / 2006

Informal Enterprises: Reflections on Existing Hypotheses

Vinod Anand*

Abstract

Employment is a policy priority in all the developing countries. The experience of these countries shows that as against the formal (organised) sector, informal (unorganized) sector contributes tremendously to employment generation. There may be slight inter-country variations in this context, but the overall picture remains the same, and informal (unorganised) sector contributes immensely to employment- creation in all the developing countries. This is also greatly supported both by theoretical and quantitative studies. For example, in the case of India, about 97 per cent of the total employment is generated by this sector and its various segments, and the remaining 7 per cent is generated through other sources like the formal (unorganised) sector and the public sector. Economic literature abounds in a number of hypotheses on the various facets and segments of the informal sector. Based on the experience of developing countries including that of India, this Paper attempts to reflect on a number of such hypotheses.

Keywords : Informal sector, formal sector, employment generation, developing countries.

JEL Classification :

1. Introduction

Economic literature normally distinguishes between informal (non-formal) and formal sectors, which are also termed as unorganized and organized sectors respectively. The distinction is well laid down by the International Labour Organisation (ILO) in terms of a number of parameters like, initial entry, nature of inputs/resources used, ownership, scale of operations, technique of production, skill acquirements, and nature of markets for the ultimate output (Meier, 1984; Romatet, 1983). For example,

- informal sector units invariably face lesser difficulties and hassles in terms of registration, government regulations and formalities in terms of norms, codes, and laws for initial entry as compared to those in the formal sector;
- informal sector relies more on indigenous resources/inputs as compared to formal sector which uses more of overseas resources;
- the ownership in informal sector is limited to family ownership, whereas in the formal sector it is invariably the corporate ownership;
- informal sector has small-scale operations, whereas formal sector is characterized by large-scale operations;

---

* The author is presently placed at the University of Lesotho (Southern Africa) as Professor of Economics. He was also a Fellow of the Indian Institute of Advanced Study (IIAS), Shimla from 2001 to 2003. Prior to that he was Professor and Head of Economics at the University of North West in The Republic of South Africa, and also at The University of Allahabad, India.

- the technique of production in informal sector is essentially adapted and labour-intensive, whereas it is imported and capital intensive in the case of formal sector;
- the skills used in informal sector are self-learnt and are acquired outside the formal school system, whereas formal sector is characterized by formally acquired skills; and
- informal sector is exposed to unregulated and competitive markets, whereas formal sector has protected markets in terms of tariffs, quotas, and trade licences.

Although this distinction is in relative terms, and does not categorically distinguish informal sector from its counterpart, yet it is widely accepted and effectively used by researchers.

Informal sector is categorized into a variety of sub-categories from different perspectives. One such perspective is based on the fact that this sector abounds in heterogeneity of economic activities, which are continuously undertaken by the various units within the sector. These economic activities may either relate to

- production and supply (i.e., primary producers and suppliers of goods) like, confectionaries and bakeries, sweetmeat shops, glassware, sheet metal, basket and cane ware, food products, knitted out-wear.
- or to procurement and supply (i.e., wholesalers and retailers etc. who buy in order to supply, and hence are neither primary producers nor ultimate demanders) like, groceries and general merchants, medicine stores, electrical shops, drapers.
- or to supply of services (i.e., producers and suppliers of services) like, flourmills, tailoring shops, laundries, haircutting saloons, repair shops.

Another perspective, which is based on the size of the enterprise in terms of the number of persons employed including the owner(s), categorizes the informal sector into

- Micro units, employing 10 or less than 10 persons;
- Small units, employing more than 10, but up to 50 persons;
- Medium units, employing more than 50 but up to 200 persons; and
- Large units, employing more than 200 persons.

Nooteboom, (1993) has also discussed the firm size while elaborating on the effects of firm size on transaction costs. According to him employing a definition of an informal sector unit is a vexing question. He says that " *for statistical purposes, the boundary between small and medium-sized business is variously drawn, in different countries or different contexts within a country, at 10, 20 or 50 persons engaged, at the enterprise level, and the boundary between medium size and large enterprise at*

100, 200 or 500 persons engaged. Other boundaries can also be found. Often, the boundaries are drawn differently for manufacturing and services, because a shop is sooner considered big than a factory." He further maintains that "*the universal best demarcation, regardless of context, is an ephemeral concept. The boundary one would like to draw depends on the type and purpose of the study.*" Drawing a distinction between universal and conventional demarcations, he says that "*universal demarcations are adhered to by everyone (at least within a country)*", and " *the virtue of a conventional demarcation is not that it claims to be the true one, but that it does make more or less sense*" at least within the domains of a given region.

In this categorization of firm size, one must remember the fact that informality is inversely related to size in the sense that smaller is the size the more informal is the unit and vice versa. In this sense, therefore, micro units (especially, the singly-owned units) are the most informal, and the large units are the least.

Liedholm and Mead (1999) have come out with another categorization based on the life cycle of the enterprises. They mention four major categories:

- New Starts: These are enterprises that have just entered into business and are just getting under way;
- Non-growing Enterprises: These are enterprises that have entered alright, but have not added anything to employment;
- Small Growers: These are enterprises that have survived the initial shocks, and have partially succeeded in adding to employment;
- Graduates: These are enterprises that have survived for a longer period and have added appreciably to employment in order to reach the middle ranges of the small enterprise spectrum.

Most of the categories of informal sector as described above are overlapping with each other. For example, micro units can either be manufacturing units, or they can be retailers or wholesalers, or even supplier of services. Further, manufacturing units can either be micro, small, medium, or large units.

In terms of what has been said above, it becomes clear that it is not easy to categorically define informal activities[1]. In fact, the informal sector itself is a fuzzy, blurred, and indistinct concept as has been rightly concluded by Peattie (1987). Given this lack of precision in definition at the micro level, she argues that macro-economic policies derived from contemporary conceptualizations of the informal (non-formal) sector will miserably fail to address the issues, which constrain informal activities in a variety of settings. While defining informal activities Rempel and Anand et al (1994)[2] have very clearly said that "*for the purposes of analysis it is important to define non-formal activities with constructs other than merely residual*". They maintain that

---

[1] Guerguil (1988), Pages 57-65

[2] Rempel and Anand(1994), Chapter 1.

- "*A weakness of the ILO (1972) Kenya Study is the definition of the formal enterprises with constructs drawn from the sub-discipline of industrial organization and then designating the enterprises that do not fit these constructs as the informal sector*";

- "*The literature which studies non-formal activities from a labour market perspective is now adopting a protected versus unprotected designation to define informal activities (e.g., Grootaert, 1992: 417; Mazumdar, 1989: 40–63, and Thomas, 1992: 56). Protection for formal activity employees is provided by closed-shop labour unions, legislated minimum wages, various labour laws, government as wage leader, a hierarchy of defined forms of advancement within large firms, and wage-efficiency relationships. Within this designation, non-formal activities are not protected via such means. The majority is self-employed who, by definition, will not be protected by a labour union. As Governments at all levels do not formally recognize non-formal activities, the self-employed and the employed are not subject to government labour legislation, they do not receive government services nor are they protected by labour, safety, health, and minimum wage legislation*"[3];

- "*An attempt to combine the industrial organization and labour market conceptualizations of non-formal activities has been made by Kavuluvulu (1990: 4) according to whom the term informal sector refers to small-scale income generating activities that are not bound by, or do not live up to formal contractual rules or regulations that are legally recognized in a given society, and the informality of the activities extends to the way of doing things characterized by ease of entry and exit, minimal start-up costs, reliance on make-shift facilities, indigenous resources and adapted technology, reliance on family labour and acquisition of skills outside the formal education system*"; *and*

- "*In practice, government responses to informal activities tend to range from benign neglect to outright hostility. The latter reflects the peri-legal (legitimate but illegal) character of many informal activities. Thomas (1992) develops this thesis at length by devoting two chapters to the informal sector, five chapters to the irregular sector and two chapters to the criminal sector.*"

After having discussed the various aspects, difficulties, and constraints in defining the informal sector, it will be worthwhile to briefly look at its characteristics. At the outset we must remember that much or most of the informal sector is entrepreneurial rather than proletarian in composition. And, therefore, what it needs most are entrepreneurial opportunities, which will create more jobs, rather than measures designed to cause confrontation between the proletarian and entrepreneurial classes (de Soto, 1989). The available research by various analysts (Liedholm and Mead, 1988; 1999; Anand and Nur, 1985; Anand, 1990; Anand, 1994A) indicates that they do share many characteristics of informal sector, as described below:

---

[3] As labour related legislation is frequently applied to enterprises that have a certain number of employees, e.g., more than five or more than ten, empirical studies of non-formal activities typically use that cut-off point as one of the means of defining what constitutes non-formal activity.

- Informal sector enterprises, essentially micro and small with not more than 50 workers engaged in manufacturing activities like light consumer goods, constitute a significant component of the industrial sector in most of the developing countries. Most of these enterprises have fewer than 10 persons, and they account for more than 50 percent of the total industrial employment in these countries. These units are generally engaged in the production of goods like, clothing, furniture, and food and beverages, and in repair work like, automotive, electrical, and bicycle, blacksmithing, and light engineering.

- Most of these enterprises are located in rural areas (that is areas having fewer than 20,000 inhabitants), and their activities are rather invisible in the sense that they take place within the household or farm compound; employment generated by these rural units invariably exceeds that generated by all the urban firms.

- Most of these enterprises are not only micro units (with fewer than 10 persons), but they are less than micro in the sense that there is a plethora of singly-owned firms that employ fewer than five persons. Because of their large numbers (and hence great employment potential) on the one hand, and relatively low incomes on the other, they naturally become the focal point for the policy makers.

- In many countries significant numbers of small entrepreneurs are female.

- The largest component of small-industry labour force is that of proprietors and family workers, accounting, on an average, for over 50 percent of small-enterprise employment. Apprenticeship labour is also important in certain countries, especially in Africa, and that too in the West Africa region.[4] On the other hand, hired workers form the smallest segment of the sector, and are generally found in the more modern type of enterprises operating on a larger scale like, bricks and tiles, baking, repairs, and metalworking.

- In most of the countries small-scale enterprise sector is characterized by a relatively narrow range of output and short period of production.

- In a number of countries a person, on an average, does not work full time (2,400 hours annually). It is interesting to see that, according to the ILO (1984) study of certain selected developing countries, the annual mean person-hours ranged from a low of 1,164 in Sierra Leone to a high of 2,514 in Egypt.

---

[4] For this see the Papers on "Apprenticeship Rules Among the Akwete Igbo of Nigeria and the Baule of the Ivory Coast", and "Hausa Weaving and Surviving amid the Paradoxes" in Coy (1989).

- Quite a number of persons working in small-scale enterprises also work part time in other activities, essentially agriculture. In real practice, it is seen that farm and non-farm employment often move in opposite directions over the year depending on the requirements of these sectors, and they are thus quite complementary in terms of providing employment opportunities to the rural people.

- The initial investment of most of the small-scale enterprises is modest. More than 50 percent of the total fixed cost is generally incurred on buildings and fixed infra-structure (including furniture, fixtures, and other fittings), and capital and other equipments account for about 40 percent of the fixed cost. A reasonably high percentage of the total variable cost accounts for raw materials, next comes labour, and smallest share goes to other inputs.

- Majority of funds for initial entry for an enterprise come from personal savings, relatives, and retained earnings. The role of the formal credit institutions in this context is indeed minimal.

- There is sufficient evidence to show that in recent decades the small-scale activity in most of the developing countries is on the rise. In fact, it has been growing at a faster pace even than the large set-ups in a few countries.

On the basis of the characteristics that have been mentioned above, it can safely be concluded that the small-scale enterprises are extremely diverse and heterogeneous. (Liedholm and Mead, 1999). This character has two important implications. The first implication is that different types of enterprises have different contributions to make to the process of economic growth and development, and the other is that they face different sets of problems and constraints. Both of these implications are of crucial importance for policy makers and the support agencies in the sense that they have to come out with different policy packages and support programmes for different sets of enterprises.

Existing Hypotheses5

We describe below the validated hypotheses on the various aspects of the informal sector enterprises in various countries:

1. Most of the developing countries face the problems of poverty, employment-creation, and income distribution, and coupled with that is the severe constraint of a highly limited scope for generating major increases in formal sector employment with the result that the rate of employment creation by the formal sector fails to keep pace

---

5 For this Section see Anand (1994B).

with the high and continuously rising rates of unemployment. And, hence reliance is placed on informal sector which develops on its own and without the support of public authorities, and quite often with their disapproval [{Page, 1979}; {ILO, 1972} as cited in {Meier, 1984}}.

2. Informal sector enterprises have both forward and backward linkages[6], the former are generated by the demand-side factors (stemming from either retailers, or wholesalers, or even consumers), and the latter by the supply-side factors (in respect of raw materials and other inputs needed by the enterprises). It is through these linkages that the informal sector enterprises interact with the rest of the economy, and they operate as participants in market systems. Rempel, Anand and others (1994) in their study on Botswana have very well demonstrated the direction of these linkages only in terms of micro enterprise responses to questions on their customers and their sources of inputs, and not in terms of the magnitude of these linkages in nominal terms. However, there is a good amount of controversy with respect to the issue of linkages, especially the inter-sectoral linkages within the economy. Some believe that informal activities have autonomous entity, and, as such, they do not generate significant inter-sectoral linkages (Sethuraman, 1976). On the other hand, there are others who strongly believe that informal activities have tentacles into other sectors of the economy in the form of backward and forward linkages (Davies, 1978; Webb, 1974). Beyond that, Dasgupta (1992) draws our attention to the analysis of backward linkages as linked with the analysis of the different survival strategies and also with the socio-demographic composition of the sector. He identifies four types of linkages in respect of petty trading in Calcutta: 1) traders who purchase bulk of their goods and services on credit; 2) traders receiving a commission on sale; 3) traders who purchase the bulk of their goods in cash; and 4) producer- sellers. Thomas (1992) sums up the relationships between sectors as follows: 1) in trading, backward linkages occur when informal activity operators obtain their supplies from the formal sector, e.g., as scavengers and by collecting industrial and consumer waste products that can be recycled as raw materials for the formal sector, hence the linkage in reverse; and 2) forward linkages occur as some informal food shops supply food to both informal and formal activity employees. It has also been experienced (Stren et al, 1992) that these linkages generate employment-creation and income-distribution effects, and, as such, they are of utmost importance if systematic increases in productivity are to be realized. Liedholm and Mead (1999) have very interesting findings in respect of Malawi, Swaziland, and Zimbabwe.[7] These relate to patterns of purchasing of inputs as well as marketing of the products. It has been found that a) over 96 per cent of the enterprises sold primarily

---

[6] Hirschman (1958) has developed a concept of linkages defining the links among enterprises in a manner that can be taken into account in formulating a development strategy.

[7] Liedholm and Mead (1999), Pages 54-58.

to individuals; and selling either to traders or to other manufacturers was relatively rare; b) about a third of the enterprises make or gather their own inputs without any reliance on the market; c) substantially higher percentage of those enterprises that sell to traders have grown in the past, and their average growth rates have been substantially higher compared to those producers that sell directly to individuals; d) there are a number of institutional arrangements which can help MSEs to move beyond a direct sale of their output to final consumers. These are: subcontracting (a contract to do part of the work that has been given to another firm), franchising (formal permission given by a company to produce or sell its brand products etc. in a particular area), flexible specialization (involving market-based relationships between independent enterprises, each specializing in different steps in the production process), and a consortium of enterprises (cooperating only to sell their products in a particular market).[8]

3. As we have already said earlier, according to the ILO study (1972) as published in Meier (1984), informal sector enterprises are characterized by ease of entry, reliance on indigenous resources, family ownership of enterprises, small scale operations, labour-intensive and adapted technology, skills acquired outside the formal school system, and unregulated and competitive markets.

4. Relative to large-scale formal sector enterprises, micro- enterprises absorb smaller shares of total investment, and account for an appreciably larger share of recorded industrial employment.[9]

5. Despite debilitating restrictions frequently imposed by authorities, micro-enterprises provide employment that can be deemed to be both efficient and profitable.

6. It is true that quite a number of micro, and even small units, fail to survive because of many factors essentially, their built-in risk aversion and the lack of sufficient demand for their products, and also because of government apathy towards their support. But this should not be misconstrued to conclude that there are no units that, despite the given 'exogenous' inhibiting factors, with time, move towards larger units. There is no lack of examples to substantiate this view. Many case studies of indigenous medium and large firms have shown that a number of them have achieved their current market position through an expansion of previously micro and small firms. There are many examples in support of this hypothesis. Some of these are:

(a) One of the famous brand names in footwear in India, called Woodland, launched formally in 1993, and living up to international quality shoes at affordable prices, and leading the market allover, has been the outcome of a small-little shoe shop in Ambala (Punjab, India) way back in 1954. Presently, it is one of the most famous large industry in shoe-making, and produces around 5000 pairs a day with a chain of 54 exclusive stores, and more than 2000 multi-brand outlets, in the whole country. It has thus been a long way from the early start in early fifties (Kanungo, 2001A);

---

[8] For a fuller discussion on these, see Mead (1994).

[9] For evidence on this hypothesis and on hypotheses 6 to 9 see Page (1979); Anderson (1982); Anderson and Khambata (1981); World Bank (1978a); World Bank (1978b); Hoselitz (1959); Staley and Morse (1965); Dhar and Lydall (1961); Berna (1960); Kilby (1962); Aubrey (1951).

(b) A small stationery shop in Delhi's Sadar Bazar, opened in 1960s, started manufacturing a pen with a brand name of Luxor, and got converted to Luxor Writing Instruments Ltd. The brand is presently the market leader in the Rs.1250 crore Indian writing instruments industry. Its annual turnover is around Rs.150 crore, and it exports to over 51 countries. The dream of a small stationery shop thus became a reality in 1996 (Kanungo, 2001B);

(c) The story of Omega begins in 1848 with a clock assembly workshop in La Chaux-de-Fonds (Switzerland). In 1979, it switched to manufacturing watches, and by 1889, it became the largest Swiss watch company with 600 employees and a production of 100,000 pieces. It went on expanding with time, and the Omega collection became one of the best in the world. Omega came to India 1n 1997 (Kanungo, 2001C);

(d) The company producing the Teacher's brand of Scotch whisky started from a small cottage industry in Scotland, and has now become the largest selling whisky in the world (Kanungo, 2001D);

(e) The history of the Goodyear tyres is also interesting in the sense that the patent of the vulcanized rubber was sold by Charles Goodyear to Charles and Frank Seilberling who founded the Goodyear Tyre and Rubber Company in Akron, USA. Today, this is the World's largest tyre company (Sinha, 2001);

(f) The largest selling Asian Paints in India had its origins in the 1940s, when during the Second World War, four young men had started manufacturing paint in a garage in Bombay. In 1945 it was converted into a private limited company under the name of Asian Oil and Paint Co. Pvt. Ltd. In 1965, the name was changed to Asian Paints (India) Pvt. Ltd. Right now, their brand folio contains many products, and they are the market leaders with over 26 per cent share of the total paint market in India. Their sales for the year 2000-2001 exceeded Rs 1,500 crore (Mazumdar, 2001).

Marsden (1990) also supports this hypothesis in the case of Sub Saharan Africa, where several entrepreneurs assisted by the Africa Project Development Facility (APDF) began their businesses on an informal basis, but they gradually expanded to larger units. Some of the examples[10] are:

---

[10] Marsden (1990), Pages 7-8

(a) An egg producer in Ghana started with less than US$200, three chicken pens, and 900 day-old chicks. He now employs more than 300 workers and has an annual turnover of more than US$1.5 million;

(b) A garment maker in Botswana started his business only with US$100 out of personal savings, a rented shed, and sewing machines, and just two apprentices. She now employs 65 workers and her annual sales are US$300,000;

(c) A Malawian left school at the age of 18 to work as a self-employed tobacco grader. He is now the owner and managing director of four companies engaged in tobacco growing and curing, commodity processing and exporting, property investment, and importation of machinery. Annual turnover exceeds US$1 million; and

(d) A family-owned conglomerate in Ghana, which manufactures clothing, spirits, furniture, textbooks, and other educational materials and imports vehicles and equipment, grew out of a small dry-cleaning shop. A recent World Bank study on Ghana reports that from 1983 onwards, when policy reforms had stimulated economic recovery, the average employment level in the MSEs almost doubled to 6.6 workers, and the average micro-enterprise operating before 1975 was able to grow into a modern enterprise.[11]

7. Micro enterprise sector is more labour intensive and more geographically dispersed and hence is more accessible to indigenous entrepreneurs and to the final consumers.

8. Compared to large-scale formal sector firms, micro and small-scale informal sector firms, on the one hand, have higher output-capital ratio (capital productivity i.e., output per unit of capital) and, on the other, they have higher labour-capital ratio (capital intensity i.e., absorption of labour per unit of capital). In other words, such firms are much more efficient than their counterparts in the formal sector. Efficiency is understood not only in terms of the ratios that have been mentioned above. It has many more connotations, as described by Marsden (1990). According to him efficiency of a firm is influenced by many factors including (a) ability to meet local and foreign competition and satisfy market demand effectively; (b) good judgment and sound business sense when making investments for expansion, modernization, and diversification; (c) personal characteristics, family upbringing, social and educational background, relevant to their performance as entrepreneurs; (d) recognizing collective endeavour, and improving personnel relations; (d) openness to new ideas for self-improvement through contacts with suppliers, customers, and other entrepreneurs in the same field; (e) calculated risk taking; (f) aiming for social recognition apart from maximization of profits; (g) diversification in terms of experience, resources, and

---

[11] Marsden (1990), Page 8.

comparative advantage; and (h) appreciation for supportive government, especially that which strengthens law and order and protects private property, and does not create unnecessary bottlenecks. Marsden (1990) has substantially substantiated all such connotations of efficiency in respect of micro and small enterprise sector in the African context on the basis of Africa Project Development Facility (APDF) data. Further evidence on efficiency of African entrepreneurs comes from a study of a sample of 106 modern manufacturing enterprises in Kenya.[12] According to Marsden (1990), '*the sample was classified into three groups: public enterprises, foreign private enterprises, and local private enterprises. The study found that locally owned private firms were the most efficient and profitable and the least protected (in terms of the effective rates of protection—ERP coefficients) of the three categories. The local private sector firms were efficient earners/savers of foreign exchange with low long-run domestic resource costs. Their financial rates of profit averaged 20 per cent, compared with 18 per cent for the foreign private firms and 15 per cent for the public enterprises.*'[13]

9. Products of micro and small-scale informal sector firms are inferior in quality and lower in cost as compared to those of the formal sector firms.

10. Micro firms abound in heterogeneity of small-scale manufacturing, wholesale and retailing with a bias towards the latter.[14] Anand and Nur (1985) in their study of Wad Medani in Sudan report that more than 64 per cent of the units were engaged in the distribution of goods and services, and in another study of Allahabad in Northen India, Anand (1991) reports that this figure was as high as 80 per cent. This hypothesis is also supported by Vandemoortele (1991) in the context of Sub-Saharan Africa, when he says that the activities of the majority of micro-enterprises are trade-related, while only a small percentage belongs to services and manufacturing sub sectors. A study of Lubell and Zarour (1990) on Dakar (Senegal) conducted in 1988/89 reveals that 72 per cent of the enterprises were in commerce, while the other 28 per cent were in production. This hypothesis is important in the sense that the potential for productive employment in services and manufacturing enterprises is much larger than that in trading activities. The Dakar Survey, as mentioned above, found that the average size[15] of an enterprise in commerce was only 1.1, compared to 3.0 in building and construction, 4.4 in services, and 4.5 in production activities. According to Liedholm and Mead (1999) hawkers dominate the show especially in the urban areas, and in rural areas, manufacturing units constitute substantially a higher percentage.

11. Micro-enterprises show a tendency to grow both quantitatively (in terms of number, both spatially and temporally), and qualitatively (in terms of horizontal and

---

[12] See Kenya (1987).

[13] Quoted from Marsden (1990), Page 12.

[14] For details on this issue see Tokman (1989).

[15] For average size see the Hypothesis 20 in later pages of this Section.

vertical growth through modernization and diversification). Liedholm and Mead (1988) conclude that small-scale industrial activity appears to be increasing in absolute terms in most of the developing countries, and that it has been growing at a faster rate than the large-scale industries in a few countries. They have also found that many small-scale light consumer goods enterprises, particularly tailoring, dressmaking, and furniture making, have grown quite rapidly despite their rivals in the large-scale industries. It has also been seen that a majority of the enterprises in most of the countries are of relatively recent origin. For example, Anand (1991) has shown that in Allahabad (India), a maximum number of units were added to the sector during 1970s and 1980s. Calloway (1973) and Aluko, Oguntoye and Afonja (1972) survey data on Nigeria reveal that fewer than twenty per cent of firms surveyed were operative prior to 1945. A Study on Kenya by Child (1977), and another on Tanzania by Schadler (1968) tell a similar story. In Botswana, Somolekae (1989) has shown that the majority of small businesses are new, and sixty per cent are just under 6 years of age. An ILO/JSPA (1989) Report on African employment maintains that most of the micro-enterprises essentially in trading, and services/manufacturing sub-sectors are well over 30 years of age. In respect of qualitative growth, Anand (1991) provides evidence on the fact that around 71 per cent of the units in Allahabad had shown horizontal growth in terms of stepping-up of the scale or level of production, and other related activities including sales, and around 26 per cent of the units grew vertically in the sense of adding further stages to the initial activities, and almost an equal number introduced modernization of one kind or another.

12.Micro-enterprise sector is normally dominated by singly owned units in the sense that most of them consist of only one person working alone. For example, Anand (1991) has shown that out of a total number of 5857 informal sector units in Allahabad (India), as many as 5245 (about 90 per cent) were singly owned, and the remaining 612 units (around 10 per cent) were jointly owned. A Study on Calcutta by Romatet (1983) shows similar results: singly owned units: 91 percent, jointly owned units: 11 per cent. In the Wad Medani Study in Sudan, Anand and Nur (1985) have shown that singly owned units constituted 84 per cent of the total, and the remaining 16 per cent were jointly owned. In Botswana, Somolekae (1989) reports that there were very few partnerships and the majority of small-scale businesses were singly owned. While looking at the small-scale sector in a number of developing countries, Liedholm and Mead (1988) maintain that virtually all firms are organized as sole proprietorships. According to Liedholm and Mead (1999) most of the surveys in Africa indicate that by and large the share of small enterprises (with 10 – 50 workers) in the informal sector spectrum range between 1 per cent to 3 per cent. The major share is held by micro enterprises. It is also indicated that a) singly-owned MSEs also have unpaid family workers, and both taken together constitute about three-fourth of the MSEs in the labour force, and b) trainees and apprentices hold a significant share of workers, and account for about 10 per cent of the labour force.

13. Although a majority of the owners possess some formal schooling, they primarily operate their businesses on the basis of informal and self-taught skills. Another interesting fact is that as one moves up the educational ladder (from primary upwards to university education) one finds that both the number of owners and the number of units decline sharply. Anand (1991) has shown that around 82 per cent of the 6530 owners of the 5857 units in the city of Allahabad (India) had low levels of formal education, only 15 per cent had attained university education, and only 3 per cent or so had some kind of technical education. In Wad Medani (Sudan) too, Anand and Nur (1985) provide evidence on the fact that more than 80 per cent of the 1612 owners of 1375 surveyed units possessed some formal education, merely 5 per cent could attain higher education, less than 1 per cent some technical education, and the remaining had no education at all. Similar results have been indicated by Silitshena (1989) for Botswana.

14. An exceptionally large number of enterprises are owned by male-owners, and female-owners are very few. At the same time it is seen that women dominate in informal activities as workers and not so much as owners. Their role, however, is highly crucial in the operation of certain typical vocations (like, knitting, dress making, and retail trade), which vary from country to country. Sufficient evidence on the first part of this hypothesis has been provided by Anand (1991) to show that male-owners account for more than 90 per cent of the total owners in the informal sector of Allahabad (India). The findings of certain studies on Botswana in this respect are, however, a lot different. According to these, the informal activity in Botswana is dominated more by women rather than by men, not only in the role of owners, but also as workers. For example, a study conducted by Alexander and others (1983) on behalf of the Applied Research Unit (ARU) of the Ministry of Local Government and Lands showed that more than one-half (53 per cent) of the businesses in Maun, Mochudi, Francistown, and Selibe Pikwe were owned and operated by women. In yet another study, conducted by Letswiti (1980) in five villages (Kanye, Moshupa, Ranaka, Lotlhakane, and Ntlhantlhe) of the southern district it was found that 60 per cent of the business managers were women. The Naledi Baseline Survey, as conducted by Marshall and others (1983), showed that 75 per cent of the informal sector businesses in Naledi (a suburb of Gaborone) were owned and managed by women. Liedholm and Mead (1999) also support the viewpoint that the women are both the owners and the workers in the majority of the micro and small enterprises in most countries. They further say that micro and small enterprises headed by women are more likely than their male counterparts to operate from the home, and '*since it is the home-based MSEs that tend*

to be overlooked, both in assistance projects and in official statistics, women owners are particularly likely to be invisible entrepreneurs."[16]

15. A good number of owners/workers are migrants. Evidence on this issue is provided by a number of studies. In his survey on Calcutta's informal sector Romatet (1983) points out that informal sector is characterized by the mobility of its labour, and it is through this sector that employment-seekers immigrate to the city, of course with the help of their friends and relatives who are already settled there. Informal sector, thus, becomes both the consequence and cause of migration. Romatet is quiet on the migration of owners, but Anand (1991) in his study on Allahabad (India) provides sufficient information on migrant-owners and concludes that more than 30 per cent of the owners in a total of 6,530 are migrants to Allahabad city from elsewhere in the district. He also talks of two other categories of migrant-owners: those who migrated to Allahabad from within the state, and those who migrated from outside the state. Taken together, these two accounts for about 15 per cent of the total. In their study on Wad Medani in Sudan, Anand and Nur (1985) have shown that about 45 per cent of the owners migrated to the city from other places, essentially from northern Sudan.

16. The amount of working capital used by most of the businesses is modest as is their initial investment. Evidence on this aspect is provided by Liedholm and Mead (1988) who have worked on data for Jamaica and Honduras, as collected by Michigan State University Country Studies, on data on Egypt, as collected by Davies et al (1985), and on data in respect of Sierra Leone, as collected by Liedholm and Chuta (1976). According to them the average figure of the initial capital stock for the small-scale enterprises in Jamaica for the year 1979 comes out to be US $1,140, in Honduras it is shown to be US $ 354 for the year 1980, and for Sierra Leone it is worked out to be only US $49 for the year 1975. As against this, the amount of working capital used in these three countries in the given years has been US $1,217, US $81, and US $50 respectively.

17. The major part of the total fixed cost is incurred on buildings, premises, and fixed infrastructure as against capital-equipments. Anand (1991) has shown this to be true for Allahabad, where cost on buildings and fixed infra-structure accounts for about 63 per cent of the total, while tools and capital equipments account for around 35 per cent of the total, and the rest, about 2 per cent, is incurred on land. The case of Wad Medani (Anand and Nur, 1985) presents almost a similar picture: the building cost is the highest (around 57 per cent), the cost on capital equipments is around 40 per cent, and the rest, around 3 per cent, is incurred on other items. There is another important finding about depreciation in both of these studies; it varies between 1.6 per

---

[16] Lidholm and Mead (1999), Page 5.

cent to 1.9 per cent of the total fixed cost for Allahabad, and between 1.1 per cent to 16.19 per cent for Wad Medani.

18. The major component of the total variable cost is normally incurred on raw materials (including intermediate inputs). Evidence, as drawn from Allahabad (Anand,1991), amply supports this: cost on raw materials accounts for 62 per cent of the total, labour cost accounts for 30 per cent, and the remaining about 8 per cent is incurred on other inputs like, power, water, telephone, and transport. The study on Wad Medani (Anand and Nur, 1985) also bears this out: 67 per cent of the cost is incurred on raw materials, 18 per cent on labour, and 15 per cent on other inputs.

19. There is a presence of substantially high fixed cost and an extremely low variable cost. Assuming that fixed cost implies generation of capacity, and variable cost implies repeated use, the wide disparity between fixed cost and variable cost among the small businesses connotes, apart from other things, under-utilization of capacity. Under-utilization of capacity can be due to both supply and demand side difficulties. According to Steel (1977). Child (1977), and the World Bank (1978), low levels of capacity utilization in the small enterprise sector occurred in Ghana, Kenya, and Zaire respectively because of the supply-side constraints in the form of shortages of imported intermediate inputs. On the other hand, there is some limited evidence provided by Page (1974) to show that in Ghana's furniture manufacturing industry smaller firms had captured an increasing share of declining market and had demonstrated higher utilization rates. Anand (1991) also provides sufficient evidence to show that small enterprise sector of Allahabad exhibits low levels of capacity utilization, the sources of excess capacity being principally supply determined. The under-utilization of capacity reduces both the labour-capital ratio (capital intensity), and the output-capital ratio (capital productivity), and, therefore, in terms of the Hypothesis 8 above, the small-scale sector becomes economically inefficient. But for a significant range of products and in a number of countries, where small-scale firms are not inhibited by demand and supply constraints, it has been shown that small-scale industry is economically efficient as compared to its larger counterparts. According to Liedholm and Mead (1988), "the labour-capital and output-capital ratios are the economic efficiency measures most frequently used in empirical studies concerned with development. These partial efficiency measures are based on the assumption that labour is abundant and capital is the only scarce resource. Virtually all the aggregate and most industry studies reveal that small-scale industries generate more employment per unit of scarce capital than their large-scale counterparts. The available evidence on relative capital productivity is somewhat limited and more mixed. Yet in the majority of countries where such comparisons have been made, the output per unit of capital among small producers is found to exceed that generated by large industry." In this context some of the studies have also used the analytically more accurate comprehensive

measure of economic efficiency where all the scarce resources are considered and evaluated at shadow or social prices. But the findings of these studies are rather mixed. Liedholm and Mead (1988) have used a social benefit-cost analysis to compare the relative efficiency of small and large firms in Sierra Leone, Honduras, and Jamaica, and conclude that in 10 to 12 specific industrial groups, the social benefit-cost ratios of small firms were greater than the comparable ratios for the large-scale firms in those particular industries.

20. Micro-enterprise sector creates substantial wage and self-employment. This is supported by a number of standard arguments. The support is also provided by a number of empirical studies. Talking in terms of wage-employment, Anand (1991) has shown that each unit in the informal sector of Allahabad employs on an average 1.6 persons. In a survey on Ahmedabad informal sector industries in India, Papola (1981) has worked out an average size of 2.8 persons per unit, and Romatet (1983) comes out with an average size of 6.2 persons for Pilkhana area of Howrah District in West Bengal. It has also been observed that the incidence of informality becomes high where firms size is small, and vice versa. Comparing the relative share of wage and self-employment in the total employment scenario in the informal sector of Allahabad, Anand (1991) concludes that the sector is dominated by wage earners (including apprentice labour) who constitute around 59 per cent of the total employment and the remaining around 41 per cent are self-employed. According to Liedholm and Mead (1988), "*proprietors and family workers generally form the largest component of the small-industry labor force, accounting on average, for over 50 per cent of small-enterprise employment. Apprenticement labor, however, is also important in some areas, particularly West Africa.*" There is also a belief that unemployed youth can easily be settled into self-employment in the informal sector, but much depends on the type of activity. If it is commerce (distributive services), access in terms of the required attributes for self-employment is much easy as compared to production activities. In the context of Sub-Saharan Africa, Vandemoortele (1991) maintains that it takes about 15 years for a school leaver, joining the labour market, to become a micro-entrepreneur in services or manufacturing sector. This period is essentially filled with apprenticeship, formal and informal sector wage employment, petty trading, and spells of unemployment. But this proposition that the majority of the urban labour force eventually get absorbed in self-employment is contradicted in number studies, and it is maintained that the major source of employment is not self-employment, but wage-employment. A census of Dar es Salaam (Tanzania) residents undertaken in 1984 had shown that only one-third of the employed population had a status of self-employment, and the remaining two-third were wage-employed. The 1984 population census of Ethiopia tells almost the same story for Addis Ababa, where about 26 per cent of the employed people were employers, 70 per cent were employees, and the remaining

about 4 per cent were either unpaid family workers or apprentice labour. The results of the 1985 living standard survey in Cote d'Ivorie (Ivory Coast) confirm these observations (Kozel, 1990) in the sense that around 72 per cent of the households in the city of Abidjan obtained income from wages and salaries, and 24 per cent were fully self-employed, and the remaining 4 per cent did not earn any income. Vandemoortele (1991) has depicted through pie charts the composition of the urban and rural employed workforce by employment status for Dar es Salaam, Ethiopia, and Cote d'Ivoire.[17]

We can, thus, generate another hypothesis to say that the majority of the urban labour force work for pay, and not for profit. This should not be misconstrued to mean that the owners of the informal sector units do not aim to maximize their profits. In fact, their profit rates are very high.[18]

21. Apprentice and wage labour become relatively more important components of total employment size as size of locality increases. This hypothesis is supported by Page (1979) and also by Anand (1991), who gives an example of Allahabad (India), Wad Medani (Sudan), and Howrah (India). His findings are presented in the following Table:

Table 1: Showing Percentage Composition of Total Employment

STUDY	CITY/COUNTRY	WAGE EMPLOYED	SELF EMPLOYED
Anand(1991)	Allahabad(India)	59.19	40.81
Anand & Nur (1985)	Wad Medani (Sudan)	50.19	49.81
Romatet (1983)	Howrah (India)	86	14

It is seen that wage-employment is the larger component of the total employment in all the three cities. It is also seen that in Howrah, (which is one of the most crowded cities of India), the wage earners constitute a high of 86 per cent of the total employment. On the other hand, Wad Medani, a tiny city of Sudan with a sparse population, shows a low of just about 50 per cent. Allahabad, lying between the two extremes in terms of population density, indicates a 59 per cent wage-employment component, which lies between the corresponding components of the two extreme cities as indicated in the above Table.

---

[17] Vandemoortele (1991), Page 12.

[18] See Hypothesis 26 in later pages of this Section.

22. With the exception of certain countries in Southern Africa and Latin America, female workers are almost negligible in the informal sector. A number of studies support this contention and speak of the dominance of male-earners both in the labour market and labour force. Evidence on Sierra Leone (Liedholm and Chuta, 1976), Kenya (Child, 1977), and Nigeria (Calloway, 1973) indicate that an exceptionally large number of men and few women are employed in small scale manufacturing of any type. Mazumdar's study (1979) on Bombay has found that female workers are in small minority in the Bombay labour market. Romatet (1983) also supports the same view and maintains that only 4.6 per cent of the hired workers are women in Pilkhana, Howarh in Calcutta's informal sector. Anand (1991) has evidence to show that more than 99 per cent of the workers are males in the informal sector of Allahabad. Individual enterprises, however, provide an exception. For example, Steel (1977) has found that more than 60 per cent of the labour force engaged in tailoring in his sample on Ghana consisted of women.

23. A good number of wage earners in the small-scale sector are literate. Literacy of wage earners is an issue, which appears to have escaped the attention of a number of surveys on small-scale enterprises. But we do have a few examples to support this hypothesis. For example, Mazumdar (1979) reports as high a figure as 85 per cent for literate wage earners among the survey sample of small-scale sector in Bombay city. Romatet (1983) gives a comparable figure of 78 per cent for Pilkhana, Howrah in Calcutta. Anand (1991) has worked out a figure of 67 per cent for the informal sector of Allahabad, and Anand and Nur (1985) indicate this figure to be around 78 per cent for Wad Medani in Sudan.

24. There exists a wage gap between small and large firms, though its magnitude differs from country to country. We have a number of examples to support this. Romatet (1983) reports a wage gap of the order of about 37 per cent between the informal sector industries and organized industries in Pilkhana area of Howrah district of Calcutta. Mazumdar (1979) reports a much higher gap for Bombay's small-scale sector, and infers that the mean earnings of the small-scale sector are about 47 per cent lower than those of the factory workers. Badgley (1978) and Van Dijk (1976 and 1978) report wage differentials of about 50 per cent for unskilled labour in Senegal and Upper Volta. Steel (1977) and Bayerlee and others (1979) report the presence of a similar wage gap for Ghana and Sierra Leone. Anand (1991) calculates only the average monthly wage of the 9,470 wage earners employed in the 5,857 units in the small-scale sector of Allahabad, and arrives at a figure of Rupees333.23, which is quite comparable with the corresponding figures as obtained in other studies carried out in various cities of India. For example, Romatet (1983) gives a figure of Rupees222.00 for Pilkhana, and Mazumdar (1979) reports a figure of Rupees270.30 for the small-scale industry of Bombay city. In essence, literature on small enterprises abounds in

evidence on the lower wages in the informal sector and the resulting wage-differentials between the small and large firms.

25. Micro-enterprise sector is characterized by longer working hours. Anand (1991) has computed the average working time per day of fully employed wage-earners of Allahabad city as 9.02 hours without there being a weekly off. Romatet (1983) also reports almost the same working time for the wage earners of Pilkhana, Howrah. Liedholm and Mead (1988), however, report the annual mean-person hours of employment as ranging from 1164 in Sierra Leone and 2514 in Egypt. In terms of daily mean-person hours, this works out to be 3.18 hours and 6.88 hours respectively.

26. Rate of return to operational cost in micro-enterprise sector is normally high, but there do exist inter-sectoral variations. This hypothesis is amply supported by Anand (1991), Anand and Nur (1985), Child (1977), and Liedholm and Chuta (1976). According to Anand, the rate of return ranges from more than 800 per cent in the south constituency sector to less than 450 per cent in the north and west constituency sectors of Allahabad, and the overall rate approximates to about 600 per cent. Anand and Nur have worked out a rate of return of about 538 per cent for Wad Medani in Sudan. Child reports a rate of return as high as 1000 per cent in a broadly based survey of small-scale sector in Kenya. Liedhom and Chuta report almost a similar situation for Sierra Leone and note the high variance in profit rates both within and among industries. As against these high rates, Simmons (1975) comes out with an extremely low range of rates of return (from 40.5 per cent to -3.1 per cent) for food processing industry in Northern Nigeria. Individual units within a sector can possibly show such low rates of return, but in aggregate there is a general agreement that rates of return in small-scale manufacturing are high, but extremely variable. In the context of net returns, Liedholm and Mead (1999)[19] maintain that their '*earlier studies*[20] *based on an analysis of detailed data on production costs in four developing countries suggest substantial differences in economic efficiency by enterprise size. In particular, the data indicate that net returns per hour of family labor are significantly higher for enterprises with 2-5 workers, compared to those enterprises with only one-self-employed person working alone. This increase in economic efficiency continues for the next higher size category, those with 6-9 workers; thereafter the number of observations in these data is small and the results more ambiguous.*' Similar results were found in a survey of MSEs in Kenya (Daniels et al, 1995). It can, therefore, be said that one-person enterprises generate the lowest net returns.

27. Micro-enterprise sector is invariably inhibited by a number of constraints relating essentially with infrastructure, accommodation, marketing, availability of basic

---

[19] Liedholm and Mead (1999), Page 5.

[20] Liedholm and Mead (1987).

inputs (like, labour, capital, and raw materials), finance, and also with risk aversion, and lack of additional demand. There are also difficulties imposed by government policies in terms of reforms, liberalization, on the one hand, and bureaucratic control and intervention, and related issues on the other. Sufficient evidence on the first part of this hypothesis is collected by Anand (1991) for the city of Allahabad, and also by Anand and Nur (1885) for Wad Medani in Sudan. There is an interesting evidence on finance in both of these studies, which emphasizes the fact that most of the funds for establishing and expanding small firms come from personal savings, relatives and friends, and retained earnings, and only few firms obtain funds from commercial banks, government, or from informal financial sources such as, money lenders. For example, Anand (1991) has shown that in small-scale enterprise sector of Allahabad, as many as 96 per cent of the firms used their own savings, and very few resorted to external finance. In Wad Medani study of Sudan, Anand and Nur (1985) conclude that the sector is completely self-financing, and there is a complete lack of easy credit facilities. According to Liedholm and Mead (1988), a lesser reliance on external finance results from the nascent state of financial markets in many developing countries, indicating the limited extent to which firms are directly reached by formal credit institutions. According to Marsden (1990), insufficient financing is a major constraint for most of the entrepreneurs in the micro and small enterprise sector, especially in the African context. He says that quite often quite viable projects have to wait for long time for the approval of commercial banks and other financial institutions, though the delays vary from country to country. In most of the cases, there is an insufficiency of short-run credit to meet working capital requirements. African entrepreneurs also face shortages of longer-term forms of financing. The reason for this is that most African countries lack the capital markets, where funds can possibly be raised. In fact, there is a lack of venture resources, which are quite familiar in other parts of the world. According to Hyuha et al (1993) informal credit is an important complement rather than substitute for formal credit, informal credit is less associated with time-consuming, bureaucratic lending procedures than formal sector, and informal credit is more expensive because of high interest rates and other charges than formal credit.

We have so far mentioned a set of twenty seven hypotheses about the informal sector and its various segments, essentially micro and small, which stand validated by a number of studies, and also supported by the experience of most of the developing countries. These can also be termed as conventional hypotheses. As against these we also come across a few newly emerging hypotheses on the basis of certain recent studies. We shall now briefly mention them below:

1. Number of micro and small enterprises is far larger than what is reported in most official statistics. The reason for this is that the smallest units, possibly because of their 'informality', and also because of survey difficulties, invariably escape the

attention of the concerned agencies and administrative set-ups, and as such official data only include firms that are registered or recognized otherwise by the authorities. This is greatly supported by the results of comprehensive surveys of five of the African countries (Botswana, Kenya, Malawi, Swaziland, and Zimbabwe) where the estimated total number of people engaged in micro and small enterprises is nearly twice the level of employment in registered, large-scale enterprises and in the public sector (Mead, 1994; Liedholm and Mead 1999).

2. Majority of the micro and small enterprises operate in rural areas (with 2,000-20,000 persons), and their share in urban locations (with at least 20,000 inhabitants) ranges between 25 per cent to 46 per cent. This has been greatly supported by Liedholm and Mead (1999) on the basis of a survey on some of the African countries. According to them, '*MSEs densities (number of MSEs, or number of people working in MSEs, per 1,000 people in the population) are sometimes higher and sometimes lower as one moves from urban to more rural localities.*'[21] It is further observed that manufacturing sector constitutes the largest segment of the micro and small enterprise sector in rural than in urban areas.

3. In a majority of micro and small enterprises total cost does not increase as fast as amount supplied, i.e. the extra cost of supplying each additional unit falls as more is supplied. This hypothesis is supported by Reid (1993) on the basis of a unique and extensive database of small businesses, containing over 40,000 data points, in Scotland in the 1980s. The results were essentially based on qualitative evidence, and it was found that 55 per cent of the units (including manufacturers, services, construction, and transport) depicted falling average variable and falling average total costs at all output levels. On the other hand, 11 per cent showed that total cost increases in line with amount supplied (indicating a linear cost curve), 15 per cent showed that at first total cost does not increase as fast as supply, but then it increases faster than supply (indicating a falling reversed-S cost curve), 14 per cent showed that total cost increases in line with supply until the maximum possible supply (full capacity) is reached, and after this point, the extra cost of supplying another unit rises sharply (indicating a linear cost curve with tight capacity). The remaining 5 per cent units did not at all respond to this query on cost curves and scale economies.[22]

4. In the context of informal sector there exists a correlation between the rates of growth of wage employment and of gross domestic product (GDP). In other words, GDP grows fastest in countries with rapid increases in wage employment This has been substantiated by Marsden (1990)[23] in the case of six African countries (Botswana, Cote d'Ivoire, Ghana, Kenya, Malawi, and Tanzania), and also in countries around

---

[21] Liedholm and Mead (1999), Page 41.

[22] Reid (1993), Chapter 4, Pages 58-61.

[23] Marsden (1990), Table 1 and Table 2, Pages 9, 10.

the world. He contends that in Botswana '*the number of wage employees rose by 8.8 per cent annually from 1965 to 1980, their share of the total labor force increased from 8 to 29 per cent, and GDP soared by 14.2 per cent annually,*' and '*the boom in wage employment and GDP has continued since 1980.*' He further maintains that '*increases in the share of wage employment were also associated with brisk GDP growth in Cote d'Ivoire, Kenya, and Malawi,*' and '*in Ghana and Tanzania, however, low rates of growth of modern sector employment coincided with slow GDP growth and a predominance of government in the economy.*' Looking at the cross-sectional data of GNP per capita and wage employees of various countries like, Burudi, India, Indonesia, Philippines, Peru, Mexico, Korea, Portugal, Ireland, Netherlands, Sweden, and U.S., it has been found that there exists a positive correlation between the share of wage employees in the total labour force and per capita GNP. Marsden (1990) says that the data on these two variables '*reflect a well-documented tendency for incomes to rise as labor moves out of agriculture (where the proportion of self-employment and family labor is highest) into industry and services and as the average size of enterprise increases to take advantage of economies of scale.*' Anderson (1982) has also examined this evidence and concludes that '*employment in household manufacturing was seen to decline first in relative and then in absolute terms as industrialization proceeds.*'

5. The informal sector is not a source of economic dynamism. This is well supported by an International Labor Organization (ILO) review of studies of the informal sector in Africa carried out by the ILO's Jobs and Skills Program for Africa (JASPA) team in many countries over several years.[24] The ILO found that '*working conditions in the sector are extremely poor*', and '*the traditional apprenticeship system used in the informal sector*' does not in any way lead to more efficient organization in the enterprises. The report further said that '*training is imparted on the job through repetitive practices by the apprentices of activities carried out by masters. Due to the heavy involvement of masters in production, time spent in specifically training apprentices is extremely low and whatever training that is provided does not involve any theoretical backup on the technologies of the machines in use and repair. The training provided does not take into account the issue of obsolescence and the increasing sophistication of machines.*' It also says that the typical manager in the informal sector '*cannot perform a correct cost pricing exercise*' and that only a few of informal enterprises keep records of any kind. The report concludes to say that a) one of the serious fallouts of the weak management '*is the extremely low life expectancy of the informal enterprises which has remained as low as five years on the average during the past decade;*' b) '*the lack of accumulation of experience in specific activity has kept the quality of the products at relatively low levels;*' and c) there has been a '

---

[24] See ILO/JASPA African Employment Report, 1989.

declining share of the informal sector in the total urban labor force in African countries.'

Political Economy of Informal Sector

In the present day context it is of much relevance to regard every economic issue or formulation as determined or influenced by what we term in the 'New Political Economy' as the 'Role of the State' in terms of its being treated as an 'endogenous' rather than an 'exogenous' variable unlike the constructs of earlier formulations within the restricted domains of conventional economics (Anand, 1996). There are other interesting aspects of the political economy of small business that relate to the ethics of competition amongst small firms and their perception of enterprise, as has been underlined by Reid (1993). The main input to small business ethics is normally drawn from the competitive strategy, competitive advantage, and defensive strategy (incorporating means for consolidating and defending an acquired market niche) of small business firms. It has been found both empirically and otherwise that

- much of the conduct of owners-managers of small firms is governed by self-interest, and not so much by any moral ethics and fair play in the sense that (a) they have no ethical qualms about seeking to establish a competitive niche by using an aggressive strategy like, pushing costs below that of rivals, or extolling the virtues of their products vis a vis those of their rivals, and (b) they are uneasy about methods or strategies (like, entry deterring/preventing prices) which they themselves would not wish to have been deployed against them when starting out in business;
- The notion of good competitor has been seen to be largely an economic, rather than an ethical concept;
- The business language used by the owners-managers normally has ethical overtones, but, by and large, it is all figurative, and lacks morality, rights, fairness, or justice; and
- The owners-managers, the basic role players in the various market games they are involved in, have no place in moral order because of the fact *that "the moral order is beyond the market, and the players put aside mere games when they participate in moral order."*[25]

Apart from what we have said about business ethics, perceptions of the business enterprise culture are also important. Enterprise culture can be defined as the pattern of attitudes and social norms likely to foster the smooth working of a free enterprise system. In fact, these perceptions relate to what the owners-managers (entrepreneurs) think of the enterprise they are linked with, and what exactly they expect form it. These perceptions affect motivation and judgment, and have a strong bearing on practical decisions taken both by entrepreneurs and by policy makers. These perceptions can either be analyzed scientifically (formally) or otherwise (less formally). For example, if we analyze and assess perceptions through quantitative evidence to find as to how the market share of micro and small firms influence their probability of survival, it will be more scientific than otherwise. But, on the other hand, if we do this exercise just on the feedback basis, it will perhaps be less scientific.

---

[25] Reid (1993), Chapter 12, Pages 211-15.

There is no doubt about the fact that enterprise culture is a must for the development of an enterprise within a given framework. There is enough evidence to show that there is a severe lack of enterprise culture in the case of firms that are prone to bureaucratic control and procedures, and which are tightly monitored and controlled by various kinds of mechanisms, including planning and state-level interventions. On the other hand, very micro and very small firms that are close to financial inception do exhibit the required enterprise culture in their own way.

Perceptions of enterprise culture have two dimensions. First, when perceptions are treated exogenously, as something apart from the entrepreneur himself, and the entrepreneur comments merely as an observer/ outsider on the business milieu. Second, perceptions are treated as endogenous parameters, as part of the entrepreneur himself, and he comments not like an observer, but as an active role player in the day-to-day business activities of his venture. The distinction is similar to one where a keen cricketer is interviewed just like a spectator, and when he is interviewed like a full-fledged player on the field. Thus, perceptions of enterprise culture can either be profiled on the basis of observance or on the basis of participation, or both.

It has been observed that in most of the countries entrepreneurs are not attuned to enterprise culture of any sort even if they are successful and highly enthusiastic about their enterprise. They may raise their voice for developing the enterprise culture, but they may not be willing to contribute directly for its formation.

Beyond perceptions, there are also many objective measures for developing the required kind of enterprise culture. Some of these measures are: access to market opportunities, access to information, and availability of informational services/net work, and consultancy services provided by qualified and trained specialists in the various fields of business ventures.

We can, therefore, conclude this paper by saying that the political economy of the informal sector, centering around the role of the State, the business ethics, the business morals, and the perceptions of the entrepreneurs for developing business culture, inhibit, to a large extent, the existence, growth, and even the survival of this sector, and contribute to negative incentive effects. Reid (1993) has very well reported on the evidence in this context on the basis of seventeen case studies out of seventy-three small business enterprises in his sample of Scotland.

References

Alexander, E.J., et al, (1983). "Informal Sector Business in Four Botswana Communities", Ministry of Local Government and Lands, Government of Botswana, Gaborone.

Aluko, S.A., O.A., Oguntoye, and Y.A.O. Afonja., (1972). Small Scale Industries: Western State of Nigeria, Ile-Ife, Industrial Research Unit, University of Ife.

Anand, Vinod and Nur, El-Tahir Mohammad, (1985). The Role of Small Enterprises in Developing Countries: A Case Study of Wad Medani in Sudan, Asian Economic Review, Vol. 27, No.1/2, April /August, Mimeographed.

Anand, Vinod K., (1990). "The Role of Non-Directory Enterprises in the Local Economy: A Case Study of Allahabad". Report prepared for the University Grants Commission, New Delhi.

Anand, Vinod K., (1991). Small Enterprises and Urban Economy, Allahabad: Chugh Publications.

Anand, Vinod K., (1994A). "Micro and Small Enterprises in selected Urban and Semi-Urban Locations in Botswana: Basic Characteristics", Indian Journal of Economics. July 1994.

Anand, Vinod, (1994B). "Informal Sector in Developing Countries: Certain Crucial Hypotheses", The Indian Economic Journal, Vol. 42, No. 2, October- December, University of Bombay, India

Anand, Vinod K., (1996). " On New Political Economy: An Overview", The Indian Journal of Economics, Vol. LXXVII, Part II, No. 305, October, University of Allahabad, India.

Anderson, Dennis, (1982), "Small Industry in Developing Countries: Some Issues", World Bank Staff Working Paper No. 518, Washington D.C.

Anderson, Dennis and Khambata, Farida (1981). "Small Enterprises and Development Policy in the Philippines: A Case Study", World Bank Staff Working Paper No. 468(July) Washington, D.C.

Aubrey, H.G., (1951). Small Industry in Economic Development, Social Research, September.

Badgley, Susanna M., (1978). Assistance for Artisan in Senegal, Washington D.C.: World Bank, West Africa Projects Department.

Bayerlee, Derek, Carl K. Eicher, Carl Liedholm, and Dunstan S.C. Spencer., (1979). Employment Output Conflicts, Factor Price Distortions and Choice of Technique: Empirical Results from Sierra Leone, East Lansing MI: Michigan State University, Department of Agricultural Economics, African Rural Economy Program Working Paper.

Berna, J.J., (1960). Industrial Entrepreneurship in Madras State, Stanford Research Institute, California, and Asia Printing House, New Delhi.

Calloway, Archibald, (1973). "Nigerian Enterprise and Employment of Youth", Ibadan: Nigerian Institute of Social and Economic Research.

Child, Frank C., (1977). "Small Scale Rural Industry in Kenya"., Occasional Paper No. 17, African Studies Centre, UCLA, Los Angeles.

Coy, Michael W.ed., (1989). Apprenticeship: From Theory to Method and Back Again, SUNY Series in Anthropology of Work, Albany: State University of New York Press.

Daniels, Lisa, Mead, D.C., and Musinga, M., (1995). "Employment and Income in Micro and Small Scale Enterprises in Kenya: Results from a 1995 Survey", GEMINI Technical Report No. 92.

Dasgupta, N., (1992). Linkages, Heterogeneity and Income Determinants in Petty Trading: the Case of Calcutta, World Development. 20:10

Davies, R. (1978). The Informal Sector: A Solution to Unemployment? Catholic Institute for International Relations, London.

Davies, Stephen, et al (1985). Small Enterprises in Egypt: A Study of Two Governorates, Michigan State University International Development Paper No.16.

de Soto, Hernando (1989). "Constraints on People: The Origins of Underground Economies and Limits to Their Growth". Mimeo.

Dhar, P.N., and Lydall, H.F., (1961). The Role of Small Enterprises in Indian Economic Development, Asia Publishing House, New York.

Guerguil, Martine (1988). "Some Thoughts on the Definition of the Informal Sector", Cepal Review 0 (35), August.

Hirschman, A.O., (1958). The Strategy of Economic Development, Yale University Press, New Haven.

Hoselitz, B.F., (1959). Small Industry in Underdeveloped Countries, Journal of Economic History, Vol. 19. Reprinted in Development Economics and Policy, Ian Livingston (Ed.), Allen and Unwin, 1980.

Hyuha, M, Ndanshau, M.O., Kipokola, J.P., (1993). Scope, Structure and Policy Implications of Informal Financial Markets in Tanzania, African Economic Research Consortium (AERC), AERC Research Paper 18, April.

ILO/JASPA, (1989). African Employment Report 1989, JASPA, Addis Ababa.

International Labour Office, (1972). Employment, Incomes, and Equality: A Strategy for Increasing Productive Employment in Kenya; also printed in Leading Issues in Economic Development, Gerald M. Meier, Oxford University Press, pages 183-190, 1984.

International Labour Office, (1984). Rural Small-Scale Industries and Employment in Africa and Asia, Geneva.

Kanungo, Anita., (2001A). 'It's Been A Long Walk', as reported in the Hindustan Times, 19 August, New Delhi.

Kanungo, Anita., (2001B). 'Penning History', as reported in The Hindustan Times, 7 October, New Delhi.

Kanungo, Anita., (2001C). 'The Last Word', as reported in the Hindustan Times, 4 November, New Delhi.

Kanungo, Anita., (2001D). 'Spirited Saga', as reported in the Hindustan Times, 11 November, New Delhi.

Kenya (1987). Industrial Sector Policies for Investment and Export Growth, World Bank.

Kilby, Peter, (1962). The Development of Small Industry in Eastern Nigeria, Ministry of Commerce, US Agency for International Development.

Kozel, V., (1990). "The Composition and Distribution of Income in Cote d'Ivoire." Living Standards Measurement Study, Working Paper No. 68. IBRD, Washington D.C.

Letswiti, Sheila, (1980). Entrepreneurship in Southern District of Botswana: A Sociological Survey, Partnership for Productivity, Gaborone (NGO).

Liedholm, Carl and Mead, Donald C. (1987). "Small Scale Industries in Developing Countries: Empirical Evidence and Policy Implications", MSU International Development Paper No. 9, East Lansing, MI: Department of Agricultural Economics, Michigan State University.

Liedholm, Carl and Mead, Donald C. (1999). Small Enterprises and Economic Development: The Dynamics of Micro and Small Enterprises, Routledge, London and New York.

Liedholm, Carl and Chuta, E. (1976). The Economics of Rural and Urban Small Scale Industries in Sierra Leone, East Lansing, Michigan: Michigan State University, African Economy Paper No. 14.

Liedholm, Carl and Mead, Donald C. (1988). "Small Scale Enterprises: A Profile", Economic Impact, No. 63, USIS, USA, Washington, D.C.

Lubell, H. and Zarour, C., (1990). 'Resilience Amidst Crisis: The Informal Sector of Dakar', International Labour Review 129 (3): 387-396.

Marsden, Keith., (1990). "African Entrepreneurs: Pioneers of Development", International Finance Corporation, Discussion Paper Number 9, The World Bank, Washington, D.C.

Marshall, J.C.M. et al., (1983). Naledi Economic Baseline Survey, Partnership and Productivity, Gaborone (NGO).

Mazumdar, Dipak, (1979). " Paradigms in the Study of Urban Labour Markets in LDCs: A Reassessment in the Light of an Empirical Survey in Bombay City ", World Bank Staff Working Paper No. 366 (December).

Mazumdar, Vandana (2001). 'Colourful Feat', as reported in the Hindustan Times, 25 November, New Delhi.

Mead, Donald C. (1994). "The Contribution of Small Enterprises to Employment Growth in Southern and Eastern Africa", World Development, Vol. 22, No. 12.

Meier, Gerald M., ed., (1984). Leading Issues in Economic Development, Oxford University Press, New York Page, John M (Jr.), (1974).

Economic Efficiency and Comparative Advantage in Ghana's Wooden Furniture Manufacturing Industry: USAID.

Page. John M., Jr., (1979). "Small Enterprises in African Development: A Survey",World Bank Staff Working Paper No. 363 (October).

Papola, T.S., (1981). Urban Informal Sector in a Developing Economy, Vikas Publishing House, New Delhi.

Peattie, L., (1987). "An Idea in Good Currency and How It Grew: The Informal Sector", World Development, 15:7, July.

Reid, Gavin C., (1993). Small Business Enterprise, Routledge, London and New York.

Rempel, Henry., Anand, Vinod K., Sunny, Grace., Babikanyisa, Vincent., Siphambe, Happy., Morewagae, Boitumelo S. (1994).

The Place of Non-Formal Micro- Enterprises in Botswana: A Study of Selected Urban and Semi-Urban Areas, Mimeo. Department of Economics, University of Botswana, Mimeo.

Romatet, Emmanuel. (1983). Calcutta's Informal Sector: Theory and Reality, Economic and Political Weekly, Vol. XVIII, No. 50, December 10.

Schadler, Karl., (1968). Small Scale Industries and Industrial Education in Tanzania, Munich, Weltforum, Verlag.

Sethuraman, S.V., (1976). The Urban Informal Sector; Concept, Measurement and Policy, International Labour Review. 114

Silitshena, R.M.K., (1989). 'Availability and Access to Training, Marketing and Extension for Small Scale Entrepeneurs and the Informal Sector', A Paper presented at a National Seminar on " The Informal Sector and Small Scale Enterprise Development in Botswana', organized by the National Institute of Development Research and Documentation (NIR), University of Botswana, July 27.

Simmons, Emmy B., (1975). The Small Scale Rural Food Processing Industry in Northern Nigeria, Food Research Institute Studies.

Sinha, Abhijit, (2001). 'A Radial Revolution', as reported in The Hindustan Times, 18 November, New Delhi.

Somolekae, Gloria, (1989). 'Women in the Informal Sector and Small-scale Entrepreneurship: Some Observations from a Recent Study', A paper presented at a National seminar on "The Informal Sector and Small-Scale Enterprise development in Botswana", organized by the National Institute of Development Research and Documentation (NIR), University of Botswana, July 27.

Staley, E., and Morse, R., (1965). Modern Small Scale Industry for Developing Countries, McGraw-Hill.

Steel, William F., (1977). Small Scale Employment and Production in Developing Countries: Evidence from Ghana, New York: Praeger.

Stren, Richard, et al, (1992). An Urban Problematique: The Challenge of Urbannization for Development Assistance. Discussion Paper prepared for CIDA, Centre for Urban and Community Studies, University of Toronto, Toronto.

Thomas, J.J., (1992). Informal Economic Activity, Hemel Hempstead, Harvester Wheatsheaf.

Tokman, Victor E., (1989). 'Policies for a Heterogeneous Informal Sector in Latin America', World Development, Vol. 17, No. 7.

Van Dijk, Meine Pieter., (1976). Developing the Informal Sector in Senegal: A Study of its Contexts and Potential, Dakar: ILO, Jobs and Skills Program for Africa (December).

Van Dijk, Meine Pieter., (1978). The Success of Small Entrepreneurs in the Informal Sector of Ouagadougou, Amsterdam: Virje Universiteit (February), Mimeographed.

Vandemoortele, Jan (1991). "Employment Issues in Sub-Saharan Africa, African Economic Research Consortium, Initiatives Publishers, Initiatives Ltd., Nairobi, Kenya.

Webb, Richard, (1974). Income and Employment in the Urban Traditional Sector in Peru, Mimeo. Princeton University, New Haven.

World Bank (1978). Small Scale Industry of Zaire, East Africa Projects Department, Washington, D.C.

World Bank, (1978a). Rural Enterprises and Non-Farm Employment, A World Bank Paper.

World Bank, (1978b). Employment and Development of Small Enterprises: A Sector Policy Paper.

Arthaniti 5 (1-2), 64-80 / 2006

International Trade and Personal Income Distribution in Developing Countries : Some Issues and Concerns**

Ananya G. Dastidar*

Abstract

In this paper we examine the literature which analyses how the personal distribution of income within a developing country is likely to be affected with greater exposure to international trade. We critically examine the methodology used in such studies and *emphasize* the need for a fresh approach towards this issue, especially in the context of developing countries whose structural features shape the way in which economic outcomes are affected by exogenous policy decisions.

Keywords : globalisation, personal income distribution, trade and wage inequality.

JEL Classification : F160, F190.

---

1. Introduction

A wave of globalisation is sweeping across developing countries the world over[1]. All around the developing world one observes a restructuring of economic policies, with emphasis on reduced state intervention, privatisation and increased openness, making for greater integration of the domestic with the world economy. In particular, trade liberalisation constitutes an integral part of such policies. As such, it is interesting to ask how the distribution of income will be affected in developing economies, as they dismantle barriers restricting the cross-border flow of goods and services, and trade more with the rest of the world. This paper reviews the current literature that explores how the personal distribution of income[2] within a developing country would be affected with increased exposure to international trade. We critically examine the methodology used in such studies and emphasize the need for a fresh approach towards analysing this issue, especially in the context of developing countries whose structural features shape the way in which economic outcomes are affected by exogenous policy decisions.

---

* Department of Business Economics, University of Delhi, South Campus, Benito Juarez Road, New Delhi - 110021. Email : agdastidar@gmail.com

** I would like to thank Deepak Nayyar for this paper. The usual disclaimer applies.

[1] See Nayyar (1995 and 2001) for a detailed discussion on the many facets of 'globalisation' policies.

[2] On the topic of income distributions, one could either talk about the *functional* distribution, which is the distribution of income across various economic categories, such as wages, rents, profits and so on; or about the *personal* distribution which refers to the distribution of income across households (or individuals) from all sources. The focus of this paper is on personal income distributions, on how this is likely to be affected by an increasing inflow of international trade in goods and services.

In what follows, first we describe the context in which the issue of international trade and income distribution has assumed importance in recent years (Section 1). Here we briefly touch upon the earlier literature that has evolved primarily in the context of developed countries of the North and that focuses on the functional rather than personal distribution of income. Thereafter we review studies on the issue of trade and *personal* income distributions (Section 2). These are mainly cross-country empirical studies on both developed and developing countries. We examine their methodology and underlying theoretical structure, drawing attention to a few shortcomings of this approach. We conclude by pointing out several crucial issues that should be taken into account while analysing the interrelationship between trade and distribution in the context of *developing* countries (Section 3).

Section 1 The Context

Concern over the distributional impact of dismantling trade barriers is not new. Generations of economists and policy makers have been concerned with the question of how the liberalisation of international trade policies might affect employment and the distribution of income within the liberalising country. This concern is reflected in the writings of Adam Smith as he comments on the plight of 'infant' industries in the face of trade liberalisation : "if those high duties and prohibitions were taken away all at once, cheaper foreign goods of the same kind might be poured so fast into the home market as to deprive all at once many thousands of our people of their ordinary employment and means of subsistence" (pp. 435-436; Smith, 1937). In nineteenth century England this issue was at the forefront of policy discourse, with the nationwide debate on the Repeal of the Corn Laws. In this context, English economist David Ricardo advocated the repeal of the Corn Laws and the adoption of freer trade policies for England (Ricardo, 1812). Freer trade in agricultural goods would depress farm prices relative to industrial prices. Thus income would be redistributed in favour of the industrial capitalists and away from the landlords in the countryside. Typically landlords did not undertake productive investments, unlike the industrial capitalists. As such this change in the distribution of income, Ricardo argued, would ultimately benefit England. This is one of the earliest instances of a policy debate on the issue of international trade and its impact on the distribution of income within the domestic economy.

Over the past decade there was a revival of the academic debate on the distributional implications of international trade, mainly in the context of the USA and developed countries of Western Europe (for a survey of these issues see, Burtless, 1995; Freeman, 1995; Richardson, 1995; and Wood, 1995). Throughout the decade of the 1980s, the economic position of less-skilled workers deteriorated steadily across the developed countries of the 'North'. Over the decades of the 1970s and the 1980s there was a marked increase in unemployment in the manufacturing sectors of all industrialized countries. This was accompanied by a very uneven distribution of the burden of unemployment; unskilled and uneducated workers faced a far greater risk of unemployment than skilled and more educated workers. Where labour laws and minimum wage legislations were largely absent, unskilled workers found employment elsewhere in the economy, although at much lower wages. In such economies, like the USA, unemployment figures remained low but there was a marked rise in income inequality[3]. By contrast, in Europe, where institutional arrangements prevented wages from falling sharply, countries such as France, Germany and Italy experienced little or even diminishing income inequality but higher rates of unemployment[4].

This period also saw a phenomenal increase in imports of cheap, manufactured goods from labour-abundant, developing countries of the South (such as China, Malaysia, Indonesia, Thailand, India, Bangladesh and Pakistan) to countries of the developed North. In addition, the decade of the 1980s also saw widespread technological changes (particularly computerisation) sweep across production processes in the North[5]. Both these factors, viz., imports of unskilled-labour intensive products from the South and skill-biased technological change affected adversely, the demand for unskilled labour in the North. Therefore, as economists and policy-makers attempted to explain falling living standards of the unskilled workers, they essentially debated which one of the two factors, among international trade and technological change could be held primarily responsible for these developments in the developed countries' labour markets.

In this literature, however, there is relatively little discussion on the role of the restrictive macroeconomic policies which were put in place in most countries of Western Europe in the eighties decade. As these policies sought to restrain inflation, they also tended to affect the labour market by slowing down the rate of growth of aggregate demand. It can be plausibly argued, therefore, that the restrictive policy regime also contributed to the problem of unemployment especially in the countries of Western Europe (for a discussion on these issues see e.g., UNCTAD, 1995 and Nayyar, 1996).

---

[3] In 1963 the weekly wage received by a man in the 90th percentile of the income distribution was about 2.91 times the wage of a man in the 10th percentile of the distribution. This ratio rose to 3.0 by by 1969 , to 3.47 by 1979 and to 4.42 by 1989 (Kosters, 1994). Also the real hourly wages of men with 12 years of schooling dropped by about 20 percent between 1979 and 1993 (Freeman, 1995).

[4] The rate of unemployment was 2.9 per cent in Europe in 1973. From 1983 to 1991, however, the average unemployment rate rose to 9.3 per cent (OECD Jobs Study, 1994).

[5] For instance, Berman, Bound and Griliches (1994), Berman, Machin and Bound (1996), Bound and Johnson (1992) and Krueger (1993) argue that the development and spread of new technologies which had a bias towards a more intensive use of knowledge inputs favoured, in particular, the skilled workers in the North The skill-bias in the new technologies raised the productivity (and wages) and the overall demand for skilled workers, relative to unskilled workers, thereby strengthening the relative economic position of the former.

With a fall in labour demand in general, less skilled workers would be particularly affected, as arguably, their lack of skills would limit their ability to switch between jobs.

Trade and Wage Inequality The above literature on trade and the skilled-unskilled wage structure has evolved almost entirely in the context of developed countries and its focal point is factor income distributions[6], rather than personal or household income distributions. The most commonly used theoretical framework in this literature is the 2 x 2 x 2 version of the Heckscher-Ohlin-Samuelson (HOS) theory of international trade (see Heckscher, 1949; Ohlin, 1933; and Samuelson, 1948 and 1949), and its corollary, the Stolper-Samuelson (SS) theorem, which embodies the distributional implications of a change in the trade policy regime (see Stolper and Samuelson, 1941)[7].

Empirical studies on the relation between international trade and factor incomes, primarily using data on the American economy, can be classified as follows:

(a) Some studies focus on relative price changes, as they attempt to test the empirical validity of the SS theorem (see for instance, Lawrence and Slaughter, 1993; and Sachs and Shatz, 1994), which relates changes in relative commodity prices to that in relative factor prices. Given that the skilled-unskilled wage rate has risen in the US, these studies examine whether the relative price of skill-intensive goods has also risen[8]. However, there is mixed evidence regarding the nature of change in relative goods prices. While Lawrence and Slaughter (1993) did not find evidence in support of the SS theorem, Sachs and Shatz (1994) found that relative prices did actually move in the 'right' direction and they concluded therefore that trade related factors were primarily responsible for the labour market trends observed in the US[9].

---

[6] Factor income distribution essentially refers to the distribution of total income between economic categories such as wages, profits, rents and interest. In the context of the US, economists focussed on the relative wages of skilled workers vis a vis that of unskilled workers.

[7] According to the SS theorem, under certain conditions (such as, perfect competition in the goods and factor markets; constant returns to scale technology; incomplete specialisation in production; at least as many factors of production as there are goods; and full-employment of all factors of production), a reduction in tariffs would hurt the relatively scarce factor and raise the returns to the relatively abundant factor within the economy.

[8] In the 2 x 2 version of the HOS model, with the assumption of a small open economy, it can be shown that relative factor prices change only when there is a change in relative goods prices. Given this, if the SS theorem were true (under the small open economy assumption), then if a change in relative factor prices is observed, it *must* be the case that relative goods prices have also changed.

[9] One reason for the difference in the estimates of relative price changes in the two studies was due to the fact that Sachs and Shatz excluded computer prices from their series of prices of high-skill-intensive goods. These authors argued that the massive decline in computer prices in the 1980s would tend to distort the overall pattern of price change in the US economy.

(b) Other studies focus on the relation between trade volumes and the distribution of income. The basic logic of these studies is straightforward. Imports of low-skill intensive goods are seen as substituting for domestic production. As such imports serve to *reduce* the demand for low-skilled labour (compared to import level in some base year) and thereby contribute to a reduction in the wages of low-skilled workers. Analogously, exports of skill-intensive products *raise* the demand for, and hence the wages of skilled workers. The main empirical issue therefore is to estimate the comparative effects of exports and imports on the relative demands for and relative wages of skilled vis a vis unskilled workers. This is the basic principle underlying the 'factor-content-of-trade' methodology used extensively in these empirical studies (see e.g., Borjas, Freeman and Katz, 1992 ; Murphy and Welch, 1991; Sachs and Shatz, 1994; and Wood, 1994), which broadly conclude that international trade is an important factor explaining the decline in employment of low-skilled workers in the US economy. In particular, Wood's study produces by far the largest estimates of the impact of trade on relative factor prices[10]. It is worth noting that the factor-content methodology draws upon HOS theory, insofar as the *pattern* of trade between the developed and less developed countries is concerned (see e.g., Wood, 1994). However while this approach draws a link between trade volumes and relative factor prices, within the 2 x 2 x 2 HOS framework, trade *volumes* do not play any role, rather commodity prices determine factor returns. In fact with the assumption of a small open economy, goods prices alone determine the prices of factors of production. In particular, in this model *absolute* factor endowments do not matter so far as factor price determination is concerned.

The evidence emerging from these studies is by no means conclusive and there is considerable dissent regarding the *magnitude* of the impact of trade. For instance, Krugman (1995) argues that even though trade with developing countries is likely to depress the demand for unskilled workers, yet such trade is unlikely to be a prime cause underlying the observed increase in wage inequality in the US. For trade with developing nations constitutes only a very small fraction of total trade (and an even smaller share of GDP) of the developed nations.

However, our main objective here is to examine the interrelation between trade and personal income distribution and we will now turn to these issues.

Section 2 Trade and Personal Income Distribution

In what follows, we first examine how the existing literature analyses the relation between international trade and personal income distributions, critically commenting on the basic approach adopted in these empirical studies. We argue that the framework adopted in these studies is inappropriate for developing economies and suggest the need for a fresh approach for analysing this issue in the context of developing countries like India.

---

[10] Wood (1994) actually adopts a variant of the factor-content methodology where he makes several adjustments, especially for differences in production techniques between the developed North and the developing countries of the South. His calculations indicate that the net impact of North-South trade on manufacturing employment in the US is roughly ten times what standard factor content estimates imply.

Broadly speaking there are two kinds of primarily, empirical studies relating trade openness and personal income distributions. These use cross-country datasets for econometric analyses, where typically the dependent variable is some measure of income inequality. One set of studies, like Spilimbergo et al. (1999) and Savvides (1998b), draw upon a theoretical structure developed by Bourguignon and Morrison (1990) to determine the set of independent variables for the inequality regressions. Others, like Edwards (1997) and Savvides (1998a), simply regress income inequality on measures of trade openness and macroeconomic indicators generally perceived as important determinants of the degree of income inequality within countries. The empirical specifications in these studies do not stem from any systematic theorisation as in the previous case. We discuss each of these in turn.

The Bourguignon-Morrisson Framework Bourguignon and Morrisson (1990) (BM for short) developed a formal structure for the basic idea that the personal distribution of income within an economy is dependent upon the pattern of ownership of factors of production, the economy's endowments of various factors of production and upon the structure of trade protection.

The structure of the BM model is as follows : there are *n* individuals, *m* factors and *p* sectors of production in the economy. Each sector is represented by a production function, $F_k = F(L_1, L_2, \ldots\ldots, L_m)$, where the production functions are assumed to exhibit constant returns to scale. There is perfect competition in the product and factor markets and it is assumed that factor markets clear, i.e. all factors of production are fully employed. Factor returns $(w_1, w_2, \ldots., w_m)$ are determined by the values of the marginal product of the factors in each sector.

Trade is introduced in this framework via the structure of goods' prices. With the small open economy assumption and no non-traded goods, commodity prices are determined solely by world prices[11]. In this case the structure of trade (i.e. what goods are exported and what are imported) is determined residually after taking into account the level of domestic demand. Net exports for each good is total production less domestic consumption of that good. Thus in this framework, what is important from the point of view of the impact of international trade on the distribution of income is the 'structure of protection' and *not* the 'structure of trade'.

---

[11] With non-traded goods, BM recognises that in this case, domestic demand patterns would also play a role in determining goods prices.

The structure of protection determines the prices of traded goods in the domestic market and thus the returns to factors of production and through this channel the distribution of income.

The distribution of income in this model is simply the vector Y of the total income earned by each individual in the economy, i.e. $Y = (y_1, y_2, y_3, ..., y_n)$. The ith individual's income, y_i is a weighted average of his income from the ownership of some subset of the m factors. The weights (a_{ij} / E_j), represent the share of the jth factor owned by individual i, as a proportion of the total endowment of factor j, E_j. The structure of ownership of the existing factors of production explains the personal distribution of income in such an economy.

Within this framework BM show that the distribution of personal incomes is dependent upon (a) endowments of factors of production in the economy, (b) prices of goods, (c) the structure of trade protection and (c) the structure of factor ownership.

BM carry out a cross-section analysis for the year 1970, using data on small and medium developing countries. The factor endowments they include as explanatory variables are unskilled labour, skilled labour, capital, export-specific land, non-export-specific land and mineral resources. BM assume that all small open economies face the same set of international prices and in this way avoid including the price term as a determinant of cross-country differences in income distribution. Their empirical result with respect to openness is that greater trade protection tends to worsen the income shares of the bottom forty and sixty per cent of the population.

Empirical Studies based on the BM Framework Spilimbergo et al. (1999) extend the BM framework. Treating the world as an integrated economic unit, they show that international prices are determined by world factor endowments, in exactly the same way as goods prices are determined by factor endowments in a closed economy. Since international prices determine factor prices within an open economy, they show that factor prices and hence the distribution of income in an open economy, is determined by (a) its internal endowments of factors of production , (b) world endowments of factors of production, (c) its structure of trade protection and (d) its structure of factor ownership. This extension allows Spilimbergo et al. to do away with prices and express income distribution solely as a function of factor endowments, in their empirical formulation.

The empirical studies by Spilimbergo et al. (1999) and Savvides (1998b) are mainly econometric analyses that aim to explain the variation in the personal distribution of income using panel datasets. The dependent variable in each case is a measure of income inequality and both papers use inequality measures from the Deininger-Squire (1996a) data set. The explanatory variables, drawn chiefly from the BM framework are measures of factor endowments, proxy measures for the ownership and distribution of factors of production across the population and measures of trade openness.

The studies differ from each other with respect to the measures of trade protection used and also with respect to the factors of production identified as being the primary ones whose endowment and distribution explain personal income distributions.

The factors of production considered in Spilimbergo et al. (1999) are arable land, capital, skilled and unskilled labour force[12]. They estimate the following equation:

$$\begin{aligned} Gini_{it} &= \alpha + \beta.A_{lit} + \delta.A_{kit} + \gamma.A_{sit} + \eta.O_{pit} + \lambda.A_{L}O_{pit} + \mu.A_{k}.O_{pit} \\ &\quad + \nu.A_{s}.O_{pit} + \theta.gdppc_{it} + \varphi.(gdppc)^{2}_{it} + \varepsilon_{it}, \end{aligned}$$

where, Gini = Gini coefficient; A_l = arable land per capita ; A_k = captial per worker ; A_s = endowment of skilled labour (skill intensity); O_p = measure of trade openness; and gdppc = GDP per capita.

This study uses seven alternate measures of trade openness[13] and the basic result regarding income distribution and openness is fairly robust to the choice of indices. The basic result obtained with respect to trade openness is that inequality *increases* in countries that are well endowed with skills when the economy opens. This result is in line with the empirical literature on wage inequality that finds trade openness increases the premium for skilled workers. While inequality *decreases* in countries that are relatively well endowed with capital when the economy opens. This result is interpreted as being in line with the rent-seeking literature (eg. Krueger, 1974), which argues that rents from the ownership of capital decreases when the economy opens up. For a sub-sample of developing economies the coefficient of O_p is insignificant while the coefficients of the interaction terms λ, μ, ν are significant. This indicates that trade openness per se, has no effect on the distribution of income in developing economies, apart from the effect on factor prices, which works out via factor endowments (this result is in line with Edwards, 1997 discussed below).

Savvides (1998b) considers endowments of educated labour force, natural resources and physical capital. This paper uses two alternate measures of trade protection and its empirical results are in line with the findings of Spilimbergo et al., viz. that the impact of trade openness upon income distribution will differ, depending upon the economy's endowment of skilled labour.

---

[12] For each factor of production Spilimbergo et al. construct a measure of relative scarcity in the following way : $Af_{it} = \ln (E_{fit}/E_{ft}^{*})$, where E_{fit} = per capita endowment of factor f in country i in year t and E_{ft}^{*} = world per capita effective endowment of factor f at time t, which is computed by weighting every country's endowment by its population and degree of openness (this is used as a weight to account for the fact that the endowments of close economies do not compete in the world markets with other factors).

[13] Spilimbergo et al. develop a new endowment-corrected measure of trade openness and use six existing indices of trade openness.

Other Empirical Studies on Trade and Personal Income Distribution Another class of empirical studies simply test the effect of trade protection on the distribution of income without considering factor endowments explicitly (Savvides, 1998a, Edwards, 1997). Both studies use alternate measures of trade protection and income distribution data from the Deininger-Squire data set. In Edwards (1997) the dependent variable is change in the Gini coefficient over time and this is regressed on an openness indicator, the initial level of GDP, the rate of growth, the rate of inflation, and change in the proportion of population with at least secondary schooling. Edwards focuses only on developing countries and finds no evidence linking openness or trade liberalisation to increases in inequality. Savvides (1998a) carries out a similar empirical exercise. He regresses change in the Gini coefficient on a measure of openness, the rate of growth, secondary schooling rate and includes a developing country dummy. He finds that trade liberalisation tends to increase income inequality in developing countries, but has no such effect for developed countries.

Section 3 An Appropriate Framework of Analysis for Developing Countries : A Few Comments

Two classes of studies on trade and personal income distribution have been discussed above. The second set of studies (Savvides, Edwards etc.) suffers from the drawback that their choice of determinants of income inequality is somewhat arbitrary and not directly derived from a systematic theoretical framework. The first set of studies does not have this problem. However we argue that the theoretical framework developed by BM and later by Spilimbergo et al., is not entirely suitable for analysing the problems of surplus labour developing economies. Such economies are characterized by open and disguised unemployment and display structural characteristics like the existence of an informal sector that may have an important bearing on distributional consequences of enhanced exposure to international trade.

In what follows we first analyse the implications of these two issues before setting out a few general comments on the need to develop an appropriate framework of analysis specific to developing countries.

On the Presence of Unemployment An important assumption in the BM framework, that of full employment of resources, is quite inappropriate in the context of developing countries. Typically developing countries are overpopulated and labour, especially unskilled labour is in abundant supply. Growth in formal, manufacturing sector employment in these countries is hardly sufficient to absorb the entire labour force, so these are characterized by large-scale unemployment, both open and disguised.

Existence of unemployment is likely to affect the way in which greater openness to international trade affects the distribution of income in developing countries. For, in the presence of unemployment, wages of workers would not increase by as much as under full employment. In this context a full employment model may not accurately capture the impact of greater trade openness on wages and thus on the distribution of income.

For instance, consider a labour (mainly unskilled labour) abundant developing country, with comparative advantage in a labour intensive commodity (say agriculture), where the capital intensive sector (manufacturing) is given tariff protection. This was the case in a number of developing countries which adopted ISI (import substituting industrialization) policies in the sixties and seventies decades. With trade liberalization, tariffs are reduced and the relative price of the labour intensive commodity increases. In the full employment model, with a rise in relative price, resources would be drawn towards this sector, leading to an increase in relative returns to labour, the factor used intensively in the production of this good. However, in the presence of unemployment, the same outcome need not be observed, as unemployment serves to weaken the bargaining position of labour vis a vis capital In this case an increase in demand can actually bring about matching increase in supply of labour, without an increase in wages, or with a much smaller increase in wages as compared to a situation of full employment.

In the full employment model, with an increase in demand for labour (following a rise in the price of the labour intensive commodity), wages increase. Assuming labour is the relatively abundant factor, presumably, overall income inequality would be *reduced*, in this situation. However, relaxing the full employment assumption allows for a larger range of possibilities, including the case where overall inequality may actually *increase*. For instance, in the presence of unemployment, the bargaining power of labour is weak relative to that of the owners of capital. In this case, a rise in commodity prices may actually increase profit shares (relative to the share of wages). As long as the ownership of capital (and hence profits) is concentrated among a relatively small share of the total population, this might lead to an increase in overall income inequality.

A related point is that in the BM framework, there is a straightforward relation between individuals' endowments and their levels of income. Returns to scarce resources are relatively higher, so that owners of relatively scarce resources earn relatively higher incomes as compared to those who own resources that are in relatively abundant supply. In the presence of unemployment and underemployment, this relation between endowment and income levels may be substantially weakened.

A simple example demonstrates how extensive underemployment and disguised unemployment, a characteristic feature of developing countries even today, affects the overall earnings distribution. Education is a scarce resource in developing countries, where the bulk of the population remains poor and illiterate (or poorly educated). In this context, according to the BM framework, the educated will earn far higher incomes than the illiterate. While this is true in general, this relation may not always be observed in the presence of unemployment and underemployment. To see why, consider the fact that in poor nations it is common to observe educated individuals (i.e. those who have completed high school or even those with college degrees) accepting jobs, which require little by way of educational qualifications. Examples include individuals with B.A. (or sometimes even M.A. degrees), working as delivery boys, security guards or auto rickshaw drivers. Educated individuals holding such jobs, receive much lower salaries than they would earn in a job commensurate with their skill or educational levels. In fact in these situations their earnings are hardly different from those of uneducated persons holding similar positions.

This example draws attention to the need to factor in the possibility of unemployed and underemployed resources into models of trade and income distribution for developing countries. For in such cases the conclusions of endowment based approaches like the BM framework, may be modified significantly.

On the Presence of an Informal Sector An important structural characteristic of developing economies is the existence of an informal sector, which often employs a large share of the labour force in these countries. Estimates indicate that developing countries tend to have a large informal sector that may absorb as much as up to three-fifths of total workforce. In fact there is a vast literature on the growing importance of the informal sector (especially informal services) in developing economies (see for instance, Agenor, 1996; Fields, 1990; Ihrig and Moe, 2000; ILO, 1995 and 1998 etc.). However there is relatively little analysis of the role of the informal sector in the empirical literature that explores the relation between trade liberalisation and the overall distribution of income.

However, the informal sector is likely to have important implications for income distribution mainly in the following way. In developing countries, typically the informal sector acts as a residual, absorbing the bulk of those who cannot find employment in the formal sector, at low wages. Generally labour laws, especially those relating to minimum wages, formation of trade unions, security of tenure etc. are not applicable to informal employment. Thus informal sector wages tend to be flexible and among the lowest prevalent in the economy.

Therefore it is important to examine the implications of trade openness for employment in the formal and informal sectors. For this is likely to matter from the viewpoint of income distribution, assuming there is significant wage differential between the formal and informal sectors. For instance, if employment created due to 'trade' (e.g., expansion of the exportables sector) is concentrated solely in the informal sector, the distributional outcome is likely to be quite different from the case of employment being concentrated solely in the formal sector, with labour absorption at much higher wage rates. In labour abundant developing countries, income inequality is likely to diminish in the second case, whereas in the former case the picture would be less clear. Obviously this difference in distributional outcome would arise due to the wage differential (for comparable job categories) between the formal and informal sectors.

The crucial point is, to correctly gauge the effect of greater openness on distributional outcomes, the analytical framework must explicitly allow for the presence of an informal sector.

In fact a range of possibilities open up regarding the distributional impact of trade, once an informal sector is allowed for. For instance, if, following trade liberalization, employment opportunities open up in urban centers and induces migration from rural areas, it may be the case that these rural migrants are absorbed primarily in the informal sector. In this case, their wages may not increase much compared to their earlier rural wage and thus there may not be much impact on the overall distribution of income.

Theoretical constructs based on the specific factors model (Jones, 1971) have incorporated an informal sector while analysing the likely impact of greater openness on the wage structure (see Marjit and Acharyya, 2005). However the empirical literature surveyed above does not explicitly incorporate this important structural characteristic of developing economies, a critical shortcoming of the analytical framework of these studies.

Some General Comments The literature that analyses the issue of trade liberalisation and income distribution empirically using cross-country data, ignores many aspects of the problem which are important in the context of developing economies.

There is little attempt in this literature to explore the various *channels* via which trade liberalisation policies are likely to affect the overall distribution of income. The common empirical approach adopted in each of the studies that explore the relation between international trade and the personal distribution of income is multiple regression analysis, with a measure of income distribution as the dependent variable and a measure of trade openness as one of the independent variables. Individual studies differ, both with respect to the 'other' independent variables (i.e. other than trade openness) included in the regression and the exact measure of trade openness used. While this approach indicates the sign and strength (or weakness) of the relation between trade and distribution, it fails to throw light on the specific channel(s) via which trade affects income distribution. This is one area in which further research needs to be carried out.

Finally, to analyse the distributional implications of a trade liberalization episode, attempts should also be made to incorporate the role of the State. In a number of developing countries the State has significant control on the resources via laws governing the inter-sectoral mobility of resources, minimum wage laws and other controls governing the behaviour of enterprises located primarily within the formal sector. In part economic activity within the informal sector is often the effect of efforts to circumvent existing government controls. As such in studying the distributional impact of trade liberalisation, it is important to study whether such episodes are accompanied, preceded or followed by policies of de-regulation and privatisation, involving a retreat of the State. For this is likely to affect the way in which trade liberalisation would affect the distribution of employment as between the formal and informal sectors. In turn this would shape the way in which the distribution of income is affected.

Another important reason for explicitly recognising the role of the State relates to its ability to exercise control on the composition of a country's trade basket via a host of trade protection measures. The composition of a country's trade basket, with respect to the relative shares of import-competing and non-competing goods, has an important bearing on the pattern of industrialization within the domestic country. In general trade protection serves to limit the extent of international competition faced by domestic industry. However, as trade controls are dismantled, the composition of the basket of tradables changes and domestic industry is often forced to undergo major structural changes in order to remain competitive[14]. The pattern of industrialisation, in turn, is likely to influence the quantum and quality of employment and the wage structure and via this channel, it affects distributional outcomes.

Therefore, in the current context, there is need to develop an analytical framework that would incorporate the change in the competitive environment or the nature of changes in competition policy, following an episode of trade liberalisation. For this is likely to have important, even if indirect, implications regarding the distributional consequences of greater trade openness and greater exposure of domestic industry to the pressures of international competition.

Conclusion

In this paper we reviewed the recent literature that explores the distributional consequences of trade openness. In particular we focussed on the relation between trade and personal income distribution in developing countries. We have attempted to draw attention to a number of pertinent issues that have received scant attention in the mainly empirical literature that explores how trade openness is likely to affect the distribution of income in developing countries. It is hoped that future research on this topic would try to address these issues directly or at least try to analyse how existing results and conclusions are modified once some of these concerns are addressed.

---

[14] Under the trading rules imposed by the WTO, import quotas are a thing of the past and there are clearly specified limits to the extent of tariff protection that countries can grant. In this context with trade liberalisation, it is all the more likely that the composition of imports would reflect domestic tastes and preferences and domestic industry is likely to face greater international competition.

References

Agenor, P. (1996) : "The Labour Market and Economic Adjustment", *IMF Staff Papers* 32 : 261-335.

Berman, E., J.Bound and Z. Griliches (1994) : "Changes in Demand for Skilled Labour within US Manufacturing : Evidence from the Annual Survey of Manufactures", *Quarterly Journal of Economics* 109 : 367-397 .

Berman, E., S. Machin and J. Bound (1996) : "Implications of Skill-Biased Technological Change", *mimeo*, Boston University : Boston.

Borjas, G., R. Freeman and L. Katz (1992) : "On the Labour Market Effects of Immigration and Trade", in G.Borjas and R. Freeman, eds., 'Immigration and the Workforce', University of Chicago Press : Chicago.

Bound, J. and G. Johnson (1992) : "Changes in the Structure of Wages in the 1980s : An Evaluation of Alternative Explanations", *American Economic Review* 82 : 371-392 .

Bourguignon, F. and C. Morrisson (1990) : "Income Distribution, Development and Foreign Trade", *European Economic Review* 34 : 1113-11132.

Burtless, G. (1995) : "International Trade and the Rise in Earnings Inequality", *Journal of Economic Literature* 33 : 800-816 .

Dastidar, A.G. (2004) : "Structural Change and Income Distribution in Developing Economies : Evidence from a Group of Asian and Latin American Countries", CDE Working Paper No. 121, Center for Development Economics, Delhi School of Economics : New Delhi.

Edwards, S. (1997) : "Trade Policy, Growth and Income Distribution", *American Economic Review, Papers and Proceedings* : 205-210 .

Fields, G. (1990) : "Labour Market Modeling and the Urban Informal Sector : Theory and Evidence", in D.Turnham (ed.), 'The Informal Sector and Evidence Revisited', OECD: Paris.

Freeman, R.B. (1995) : "Are Your Wages Set in Beijing ?", *Journal of Economic Perspectives* 9 : 15-32 .

Heckscher, E. (1949) : "The Effect of Foreign Trade on the Distribution of Income", in H.S. Ellis and L.A. Metzler (eds.), 'Readings in the Theory of International Trade', Blakiston : Philadelphia .

Ihrig and Moe (2000) : "The Dynamics of Informal Employment", Federal Reserve Board : International Finance Discussion Paper No. 664.

ILO (1995) : "World Employment Report", ILO : Geneva .

ILO (1998) : "World Employment Report 1998-99", ILO : Geneva .

Jones, R.W. (1971) : "A Three-Factor Model in Theory, Trade and History", in J. Bhagwati et al (eds.), 'Trade, Balance of Payments and Growth', North-Holland : Amsterdam, pp.3-21.

Kosters, M.H. (1994) : "An Overview of Changing Wage Patterns in the Labour Market", in J. Bhagwati and M.H. Kosters (eds.), 'Trade and Wages', AEI Press : Washington D.C., pp : 1-35 .

Krueger, A. B. (1993) : "How Computers have Changed the Wage Structure : Evidence from Microdata 1984-89", *Quarterly Journal of Economics* 108 : 33-60 .

Krueger, A.O. (1974) : "The Political Economy of the Rent Seeking Society", *American Economic Review* 4 : 291-303 .

Kuznets, S. (1955) : "Economic Growth and Income Inequality", *American Economic Review* 45 : 1-28.

Lawrence, R. Z. and M. J. Slaughter (1993) : "Trade and US Wages : Giant Sucking Sound or Small Hiccup ?", *Brookings Paper on Economic Activity* 2 : 161-226, Brookings Institution : Washington D.C.

Marjit, S. and R. Acharyya (2003) : "International Trade, Wage Inequality and The Developing Economy : A General Equilibrium Approach", Springer Verlag : Amsterdam.

Murphy, K.M. and F. Welch (1991) : "The Role of International Trade in Wage Differentials", in M.H.Kosters (ed.) 'Workers and their Wages : Changing Patterns in the United States', pp: 39 – 69, AEI Press : Washington D.C.

Nayyar, D. (1995) : "Globalization : The Past in our Present", Presidential Address, Indian Economic Association, reprinted in *Indian Economic Journal* 43 : 1-18 .

Nayyar, D. (1996) : "Free Trade : Why, When and For Whom", *Banca Nazionale del Lavoro Quarterly Review* XLIX : 333-350 .

Nayyar, D. (2001) : "Globalization : What Does it Mean for Development", in K.L. Jomo and S. Nagaraj (eds.), 'Globalization Versus Development', Palgrave : London

OECD (1994) : "The OECD Jobs Study : Evidence and Explanation", OECD : Paris.

Ohlin (1933) : "Interregional and International Trade", Harvard University Press : Cambridge .

Ricardo, D. (1812) : "The Principles of Political Economy and Taxation", (with an introduction by Donald Winch), 1973 : London .

Richardson, J.D. (1995) : "Income Inequality and Trade : How to Think, What to Conclude", *Journal of Economic perspectives* 9 : 33-55 .

UNCTAD (1995) : "The UNCTAD Report on Trade and Development", United Nations : Geneva.

Sachs, J. and H. Shatz (1994) : "Trade and Jobs in US Manufacturing", *Brookings Paper on Economic Activity* 1: 1-84, Brookings Institution : Washington D.C.

Samuelson, P. (1948) : "International Trade and the Equalisation of Factor Prices", *Economic Journal* 58 : 163-184 .

Samuelson, P. (1949) : "International Factor Price Equalisation Once Again", *Economic Journal* 59 : 181-196 .

Savvides, A. (1998a) : "Trade Policy and Income Inequality : New Evidence", *Economic Letters* 61 : 365-372 .

Savvides , A. (1998b) : "What Determines Income Inequality ? Evidence from The Kuznets Process", mimeo .

Smith, A. (1937) : "An Inquiry into the Nature and Causes of the Wealth of Nations", Modern Library : New York.

Spilimbergo, A., J.L. Londono and M. Szekely (1999) : "Income Distribution, Factor Endowments and Trade Openness", *Journal of Development Economics* 59 : 77-101 .

Stolper, W. and P. Samuelson (1941) : "Protection and Real Wages", *Review of Economic Studies* 9 : 58-73.

Wood, A. (1994) : "North-South Trade, Employment and Inequality : Changing Fortunes in a Skill-Driven World", Clarendon Press : Oxford .

Wood, A. (1995) : "How Trade Hurt Unskilled Workers", *Journal of Economic Perspectives* 9 : 57-80 .

Arthaniti 5 (1-2), 81-95 / 2006

Dirtier Trade for India?
The story of globalization

Sudeshna Chattopadhyay*

Abstract

With the opening up of trade a country can enjoy consumption benefits from a good without incurring the costs imposed on the local environment by the production process. Thus an explicit trade flow has an implicit pollution content. The direction of pollution flow depends on the volume and the composition of the trade matrix and the deeper root of the problem lies in the state of technological knowledge. Given the technological alternatives a developing country is more likely to specialize in environmentally intensive production processes compared to the developed countries. This propensity can be explained in terms of weaker environmental regulations. So, trade may create some apparent comparative advantage for less developed countries in pollution intensive goods. This surmise, if correct, would encourage dirty industries to migrate from 'North' to 'South'. So, trade encourages specialization in and migration of industries, which may eventually make the trade induced growth path of "South" unsustainable.

In this paper we have tried to estimate the pollution content of India's trade over the phase of liberalization and globalization. It has been observed that the pollution content of India's trade has definitely increased between 1985 and 2000, indicating a tendency towards specializing in relatively dirty industries, though no conclusive evidence regarding "migration" of dirty industries is observed.

Keywords : dirty industries, environmental regulation, apparent comparative advantage, migration, pollution control.

JEL Classification : F18, Q56

---

1. Introduction

A lot of debate is going on, on the question of the effect of trade liberalization on the environment. Generally, the developing world is facing weaker environmental regulations. This regulatory failure may lead to apparent comparative advantage in environmentally intensive product. The alternative positions in this regard are stated in terms of "pollution haven hypothesis" and "factor endowment hypothesis"[1]

---

* The author gratefully acknowledges the help, guidance and encouragement received from her Ph.d supervisor Prof.Sarmila Banerjee during every stage of this work. The comments of an anonymous referee also helped in improving the presentation of the content with appropriate clarifications. Of course, the author is solely responsible for the shortcomings.

Address for Correspondence : Flat-7, Block-II, 69L Prince Baktiar Sha Road, Calcutta -700033. e-mail: sudeshna.chattopadhyay@gmail.com

[1] The pollution haven hypothesis predicts that in the post liberalization period pollution will increase in the developing countries, as dirty industries will migrate from the developed to the developing countries. The alternative to the pollution haven hypothesis, the factor endowment hypothesis on the other hand, predict that the capital intensive industries will migrate to the developed countries from the developing countries. The factor endowment hypothesis argues that since the capital intensive heavy industrial sectors are subjected to protection in the developing economies, removal of trade barriers will encourage these sectors to expand in the developed regions. Thus the factor endowment hypothesis predicts a decrease in pollution intensity of output of the developing countries in the post liberalization period.

With trade liberalization, the differences in environmental standards between the developed and developing countries might induce "environmental load displacement". Trade enables a country to consume a good without incurring the pollution externality created during its production. So, a country importing a good effectively displaces the implicit pollution load to the exporting country. Trade liberalization might encourage less developed countries which have weaker environmental standards to specialize in pollution intensive product, resulting in a net displacement of pollution load from the developed to the developing countries. Since trade liberalization imposes environmental costs on developing economies, trade liberalization may not be always welfare improving for these countries. Hence benefits of trade liberalization must be weighed against the cost of additional environmental degradation. A number of studies have tried to measure the effect of trade liberalization on the environment (Robison 1988, Tobey1990, Low and Yeats 1992, Lucas, Wheeler and Hettige 1992, Mani and Wheeler 1999, Muradian et al 2002, Cole 2004 etc.).

The standard approach in the literature analyzing the effect of trade liberalization on the environment is to divide the effect of trade liberalization on the environment into three parts: scale effect, composition effect and technique effect (Copeland and Taylor1994 & 1995, Cole 2000 and Antweiler et al (2001)). Trade liberalization raises the level of economic activity in the economy which in turn leads to an increase in the level of pollution. This is known as scale effect. With trade liberalization there is increased production of goods in which the country has comparative advantage. This is referred to as the "composition effect". Thus for a developing country with apparent comparative advantage in pollution intensive goods composition effect is likely to lead to an increase in production of pollution intensive goods. Trade liberalization leads to an increase in the level of income in the economy which subsequently leads to the adoption of cleaner techniques. This technique effect has beneficial effect on the environment.

Indian economy is following the path of gradual liberalization since the beginning of the Rajiv Gandhi era in 1985. Before that time we had a fairly restricted economy where under the banner of self reliance a number of protectionist policies had been active. In July 1991 the liberalization and globalization policy was formally announced and since then we are trying to achieve higher and higher targets in terms of the rate of growth by increasing our degree of openness. An analysis of India's export performance over the period 1991-92 to 1999-2000 reveals that Cotton Yarn, Drugs and Pharmaceuticals and Metal Manufactures have increased their share in total exports (CMIE, July 2000). All these commodity groups are classified as dirty by Central Pollution Control Board (CPCB). At this juncture we thought it would be interesting to check whether our success in the international market is coming at the cost of higher domestic pollution or not. If trade liberalization has encouraged our country to specialize in pollution intensive industries then the benefits brought about by trade liberalization should be weighed against the cost of additional environmental degradation. Hence increased specialization in pollution intensive industries would suggest the need for efficient use of economic instruments to deal with these environmental externalities. The successful combination of trade and environment policies will not only lead to growth, but also make the process of growth sustainable in the long run.

In this paper attempt has been made to analyze the effect of trade liberalization on our trade vector. Attempt has also been made to subdivide the total effect of trade liberalization into scale, composition and technique effect. To find out whether in the post globalization era specialization in dirty goods has increased two alternative methodologies have been adopted. In the first part of the paper the Revealed Comparative Advantage (RCA) Indices for all categories of exports for the years 1985[2], 1992, 1996 & 2000 have been computed. Revealed Comparative Advantage Index of a country in a particular industry is defined as the share of that industry in the total export of the country relative to the share of that industry in total world export. If the RCA index is greater than one then the country is said to have a revealed comparative advantage in that industry. A comparison of the RCA indices in the pre and post liberalization era should help us to identify the groups of commodities towards which our comparative advantage has shifted in the post liberalization era. This should give us an idea about the "composition effect" of trade liberalization. If in the post liberalization period our comparative advantage has shifted towards dirty goods then the RCA indices of these goods should show improvements in the post liberalization period.

The second method adopted runs in terms of the estimation of total pollution content of India's trade. Any change in the structure of our exports should be reflected in the "pollution content" of our trade vector. The total pollution content of trade can be defined in terms of direct and indirect abatement cost of the exportables vis–a–vis that of the importables. An explicit trade flow has implicit pollution content. When a country imports a commodity it enjoys consumption benefits from that commodity without incurring the pollution cost generated in its production process. Thus by importing a commodity a country effectively exports or displaces the pollution load that would have been generated if the importing country had to produce the commodity itself. Similarly, when a country exports a commodity it effectively imports the pollution it had to incur to produce the commodity. In a typical developing country, according to Prebisch –Singer hypothesis the major exports are related to primary sector whereas the major imports are related to the secondary sector. If along with trade liberalization the economy grows, then it will gradually be more industrialized and the relative importance of the secondary sector activities in the export sector will go up. So, compared with the initial condition the pollution content of export would go up relative to that of imports.

---

[2] The year 1985 refers to financial year 1985-86, 1992 refers to financial year 1992-93, 1996 refers to financial year 1996-97 and the year 2000 refers to financial year 2000-2001.

When a developing country exports primary products which have very small pollution content, and imports secondary products which have very high pollution content then the economy is a net exporter of pollution. In this paper two different definitions of pollution content of trade has been introduced. The first is the total pollution content of trade. Total pollution content of trade is defined as the ratio between the total pollution content of exports to that of the total pollution content of imports. The second is referred to as the relative pollution content of trade. This is defined as the ratio of pollution content per rupee of exports to the pollution content per rupee of imports. The pollution content of India's trade matrix has been calculated for four years, viz., 1985, 1992, 1996 & 2000. The year 1985 is chosen as the benchmark. The year 1992 is the year immediately following the official announcement of liberalization whereas the years 1996 & 2000 chosen at four year intervals would show the state of pollution trading at a matured stage of globalization. Rest of the paper will be organized as follows :

Section 2 will discuss the methodology of this exercise whereas section 3 will report the major findings. The concluding section (section 4) will indicate the direction of future research.

2 Methodology

2.1 Revealed Comparative Advantage Index

Following Low (1992), the Revealed Comparative Advantage Index for commodity i for country j has been defined as:

$$RCA_{ij}=(x_{ij}/X_j)/(x_{iw}/X_w)$$

where x_{ij} = India's total export in the ith commodity, X_j = India's total export, x_{iw}= World export of commodity i, X_w= Total world exports. RCA>1 indicates a revealed comparative advantage. To compute the revealed comparative advantage indices we have used trade data in SITC categories at the 3- digit level. The data has been obtained from then UN Trade Statistics Yearbooks.

2.2 The Pollution Content of Trade

This exercise aims at finding the ratio of pollution content of India's exports to that of India's imports at four different points in time: 1985, 1992, 1996 & 2000. The pollution content of the export or the import bundle is defined as the total abatement cost implicit in the export or import bundle. When an exportable is produced direct pollution is generated by the process of production. Moreover, the used inputs, when produced, have also generated pollution externalities. So, the total pollution content (direct plus indirect) of a commodity can be estimated in terms of the total abatement cost involved in the entire process. Following Dean (1992) we refer to this externality as "abatement content" of the relevant product.

Abatement content of any commodity can be broken into two parts: (a). direct abatement content (b). indirect abatement content Direct/explicit abatement content gives a measure of the amount of pollution generated while producing the commodity itself. If $E = (E_1, E_2, \ldots E_n)'$ is the export bundle and the corresponding unit abatement cost vector is $C = (C_1\ C_2 \ldots C_n)$ then $CE' = (E_1C_1 + E_2C_2 + \ldots + E_nC_n)$ will simply give a measure of the explicit/direct abatement content of the export bundle. Direct abatement content misses out a significant portion of the abatement cost implicit in the production process of the inputs used in the production of commodities in the export bundle. Hence, to find the pollution content of the export or import bundle one must also find out the abatement content of the inputs used in producing the commodities contained in the export or import bundle. The abatement content of the inputs used in producing the export or import bundle is referred to as the indirect/implicit abatement content of the export or the import bundle. The sum of the explicit and implicit abatement content of any commodity bundle may be referred to as the total abatement content of that commodity bundle.

This concept of total pollution content can be best captured in the context of Leontief Static Open Model (LSOM) where the underlying economy is characterized by intersectoral dependence. If there are n sectors in the economy and E is the final export vector that the system wants to support, then $X = AX + E$ or, $X = (I-A)^{-1}E$ [X is (nx1) output vector needed to support E, A is the (nxn) input-output matrix)]

So the total pollution content of E will be equal to $\theta = C'X$ i.e. $C'(I-A)^{-1}E$ [C is the abatement cost vector representing the abatement costs as percentage of output of the different I-O sectors] and for import (M) would be $\delta = C'(I-A)^{-1}M$. i.e., the pollution content of trade can now be expressed as:

Total pollution Content of trade (TPC) = θ/δ

Relative pollution content of trade (RPC) = $(\theta/E_v)/(\delta/M_v)$,

(E_v - Value of total exports & M_v - Value of total imports).

1.2.1 Data Description & Data Matching

The aim of this exercise is to find out the pollution content of trade for India at four points of time 1985, 1992, 1996 & 2000. As is evident from the discussion above, data on the values of export and imports for each of these years, I-O tables for each of these years and the abatement costs of the I-O sectors are required for this purpose.

Annual export import data are reported by the UN Trade Statistics Yearbooks, in SITC category up to 5 digit level classification. Export import data has been collected for the years 1985, 1992, 1996 & 2000 at 3-digit level. For 2- digit level, the industry classification would be overly aggregative, whereas at the 4 or the 5 digit levels there exists severe missing observation problem due to small magnitude of the relevant transaction figures. For any particular category, transaction data is reported provided the figure exceeds some critical level. So, as a compromise the 3-digit level has been chosen. Input output tables have been taken from the input output transactions tables published by the Central Statistical Organization (CSO) at an interval of 5 years. Since the input-output tables for the years under consideration are not available, for finding the pollution content of trade for a particular year, the last input-output table published before that year has been used[3]. For finding the pollution content of trade in 1985 the commodity x commodity I-O table for 1983-84 has been used. For 1992, the 1989-90 I-O tables have been used. The 1993-94 I-O tables have been used for 1996.For the year 2000, the 1998-99 I-O tables have been used. Abatement cost data for the Indian industries has been obtained from Annual Survey of Industries (ASI) 1997-98, the Central Statistical Organization (CSO). The Annual Survey of Industries 1997-98 reports both running expenses and gross value of plant and machinery for pollution control for all NIC-87 categories. The ASI 97-98 also reports the gross output of all the NIC-87 categories. For the calculations, per unit abatement cost figures have been obtained by taking the ratios of the running pollution control expenses to the corresponding outputs.[4]

The rows and the columns of the CSO's, I-O (commodity x commodity) tables represent the groups of commodities which are the principal products of the I-O sectors to which they belong. For e.g., the first commodity group is Food Crops. Similarly the second commodity group is Cash crops.

The available I-O tables report the rupee value of the inter industry transactions in lakhs of rupees. i.e., if i and j are any two commodity groups in the commodity x commodity I-O table then the I-O table records the total value in lakhs of rupees of the amount of inputs of commodity group i that is used in the production of commodity group j and vice versa.

---

[3] The implicit assumption is that in the intermediate period between the publications of the two input-output tables the technology remains constant.

[4] One problem behind using constant abatement cost ratios is that in this case the per unit abatement cost is taken as fixed even though increasing average or marginal abatement cost is the usual assumption in economics. The implicit assumption behind constant unit abatement cost is that environmental standard is imposed at the unit level of production. Increase in the volume of production calls for increased abatement effort at the extensive scale but does not necessarily imply an increase in the intensity of abatement activity. Since we assume that the intensity of abatement activity is constant, per unit abatement cost is fixed. Constant abatement cost ratios have been assumed in a number of studies (see for example Robison(1988)).

The discussion in the Section 2.2 reveals that calculating the abatement cost of the export/ import bundle an input output coefficient matrix is needed i.e., the a_{ij} coefficient matrix is required. The available I-O transactions tables however record the intersectoral transactions at the aggregate level i.e., the X_{ij}s. To find the coefficient matrix from the transactions table, entry in each cell is divided by the gross output of the corresponding commodity group i.e., $a_{ij}=X_{ij}/X_j$. For example the value of transactions from sector 1 to sector 1 in the input output transactions table for 1983-84 was Rs. 328390 lakhs. While the gross value of output of sector 1 is Rs. 3492584. Hence a_{11} = 328390/ 3492584 = 0. 094. Transactions from sector 2 to sector 1 is Rs. 1364 lakhs. Hence a_{21} = 1364/ 3492584 = 0. 0004. Since both the numerator and the denominator of the expression represent rupees in lakhs, hence, the a_{ij}s are pure numbers indicating the value of ith commodity required for producing a rupee worth of the jth commodity.

The I-O table for 1983-84 is available at the 60 sector classification. The 1989-90 and 1993-94 and 1998-99 table is available at the 115 sector classification. Hence the 1989-90 and1993-94 and 1998-99 tables have been aggregated to 60 sectors following the sector descriptions provided by the CSO in their appendix IV of the I-O transactions tables for 1989-90 and 1993-94 and 1998-99.

At the next stage of the exercise the SITC trade statistics have to be matched with the production I-O data and then these I-O figures must be matched with the NIC abatement cost data.[5]

The methodology for finding the abatement cost of individual I-O categories can be illustrated with the abatement cost calculations for I-O 13 (Food Products excluding Sugar) for 1985. I-O category 13 is "Food Products excluding Sugar". It contains 13 SITC categories like meat Fresh, chilled Frozen (011), Butter (023), Shell Fish Fresh Frozen (036), Cereal etc Preparations, Veg. etc Fresh Simply Preserved (054), Fruit Nuts, Fresh Dried (057), Fruit Preserved, Preparations (058), Coffee and Substitutes (071), Tea and Mate (074), Spices (075), Feeding Stuff For Animal (081), Fixed Veg. Oil Soft (423), Fixed Veg. Oil Non Soft (424). For each of these 13 categories we know the abatement cost from the corresponding NIC categories. Now, to find out the overall abatement cost for the I-O category, the weighted average of the abatement costs of these SITC categories has been taken, with the relative trade share in value terms as relevant weights (see column 9 of table 2.1). The calculations for abatement costs for I-O 13 are shown in the table 2.1. Abatement cost ratio for I-O category 13 has been found to be equal to0.000065or 0. 0001(approx), i.e. the abatement cost of category 13 is 0.01% of the total value of output.

---

[5] See appendix for detailed description of the steps.

Table 2.1

Abatement Cost Table 1985 (An illustrative excerpt)

1	2	3	4	5	6	7	8	9
I-O Category	SITC Category	Description Of SITC Category	"Value Of Export(1985)(In Rs'000)	"Value Of Import(1985)(In Rs'000)	Total Value of trade(1985)(In Rs'000)	NIC Category	per unit abatement cost of NIC category	Calculation Of a.c. for I-O Category
	011	Meat frsh, chilld frozen	58628		58628	Slautering,Preparation &Preservtion of meat(200)	0.0004	0.000003
	023	Butter		21036	21036	Mfg of dairy prod(201)	0.0001	0.000013
	036	Shell fish frsh frozen	307215		307215	Processing,canning and preserving of Fish etc(203)	0	0.000001
	048	Cereal etc prep		22857	22857	Grain Milling(204)	0	0.000004
	054	Veg. Etc frsh, smply prsvd	55677	156386	212063	Canning&Pres.of fruits&Veg(202)	0.0003	0.000003
	057	Fruit nuts, frsh dried	222109	51483	273592	Canning&Pres.of fruits&Veg(202)	0.0003	0.000003
	058	Fruit Pres. Preparations	44114		44114	Canning&Pres.of fruits&Veg(202)	0.0003	0.000003
	071	Coffee and Subs.	225677		225677	Coffee curing,Roasting, Grindingetc(214)	0	0
	074	Tea and Mate	516711		516711	Processing and Blending of Tea(213)	0.0001	0.000006
	423	Fixed Veg Oil Soft		332626	332626	Mfg. of veg. oils and fat other than hydrogenated(211)	0	0.000009
	424	Fixed Veg Oil Nonsoft	35496	285651	321147	Mfg. of veg. oils and fat other than hydrogenated(211)	0	0.000009
13			1821597	870038	2691636			0.000065

So, the following information has been extracted from the data: (i).the input–output coefficient matrices, (ii.) The export and import figures corresponding to the different I-O sector i.e. the export /import vectors for the four years under consideration & (iii) the abatement cost vector. Now the method described in section 2.2 can be used to obtain the pollution content of trade for the four years.

2.2.2 Methodological Algorithm

Before leaving this section, for the sake of better comprehension let us repeat the steps involved in calculating the pollution content of exports and imports.

Step 1: Start with an n × n I-O matrix.

Step 2: Match the trade data (SITC categories) with I-O categories.

Step 3: Match the abatement cost data (NIC categories) with SITC categories and then using Step 2 to I-O categories.

Step 4: Find the vector representing the direct and indirect commodity requirement for producing the export (import) vector

Step 5: Find the abatement content of the exportables (importables) by multiplying the vector found in step 4 by the abatement cost vector.

Step 6: Find out the pollution content of trade following the formulae given in section 2.2

Following these six steps in the next section an attempt is made to estimate the pollution content of India's trade for four specific years: 1985, 1992, 1996 and 2000.

3. Results and Interpretations

3.1 Results

The RCA indices were computed for all categories of exports for all the years under consideration. 1985 has been taken as the base year and the RCA indices of the other three years were compared to those in 1985. It has been found that for all three years, among the categories whose RCAs showed improvement with respect to the base year, more than 60% were dirty. Between 1985 and 1992, 62% of the export categories, which showed improvement in their RCAs were dirty. Between 85 and 96, 67.9% were dirty. Between 85 and 2000 the corresponding figure was 67.2%.The commodities identified as dirty in our study are those that have been classified as orange or red category dirty industries by the Central Pollution Control Board (CPCB) and the Central Statistical Organization (CSO) of India. Table 3.1 reports RCA indices of a few selected SITC categories. A number of categories identified as dirty by CPCB and belonging to SITC categories 51 (Organic Chemicals), 53 (Dyeing, tanning & colouring materials), 59 (Chemical materials and products nes), 67 (Iron &Steel) and

69 (Manufactures of metal nes) have shown significant increase in RCAs over the period of analysis. Thus our analysis in terms of the RCA indices suggests that there has been a shift towards specialization in dirty industries.

Table : 3.1

India's Revealed Comparative Advantage Indices
(a few selected industries)

(*stands for 'insignificant transaction' for calculation of RCA index)

SITC category	1985**	1992	1996	2000
511(Hydrocarbons nes and derivatives)	--*	--	0.99	1.05
513 (Carboxylic Acid and their derivatives)	--	0.48	1.19	2.16
516 (Other Organic Chemicals)	--	1.92	3.9	5.99
531(Synt. Dye,Nat Indigo Lakes)	2.21	6.31	7.23	8.25
591(Pesticides,disinfectants)	1.13	1.83	3.34	3.9
671(Pig and Sponge iron...)	0.08	4.09	3.28	2.35
695(Tools for use in the hand or machines)	1.02	1.4	1.28	1.48
697(Houehold Equip. of base metal, nes)	0	1.97	0	5.26
699(Manufactures of base metal, nes)	0.72	--	0.95	1.58

Table : 3.2

Pollution Content of India's Trade (in Rs'000)

		1985 **	1992	1996	2000
(A)	Pollution Content of Exports	13196218	63585419	120099179	234622005.9
(B)	Pollution content of Imports	27484349	65004807	156264425	200812202.7
(A/B)	Total Pollution content Ratio	0.48	0.978	0.769	1.17
(A/Ev)/(B/Mv)	Relative Pollution content Ratio	0.835	1.13	0.948	1.43

**** 1985 refers to financial year 1985-86, 1992 refers to financial year 1992-93, 1996 refers to financial year 1996-97 and the year 2000 refers to financial year 2000-2001.**

It is interesting to note that compared to the pre-liberalization phase, in the post liberalization era the pollution content of India's trade has gone up. Impact of trade liberalization on pollution content can further be subdivided into three effects: scale effect, composition effect and technique effect. Scale effect refers to the increase in pollution content due to the increase in the export import ratio. The composition effect refers to the increase in pollution content due to a shift in the composition of the export bundle towards more polluting goods. Trade liberalization increases the level of income in the economy, which in turn results in increased demand for cleaner environment, and subsequent adoption of cleaner techniques. This is called the technique effect. Technique effect is likely to reduce the pollution content. The change in the pollution content ratios between the pre and post globalization period is likely to reflect the combined influence of all these three effects. The total pollution content ratios reflect all three effects. The relative pollution content ratios measures the pollution content of a rupee worth of exports relative to the pollution content of a rupee worth of imports. Hence, the relative pollution content ratios reflect only the composition and the technique effects. The increase in the relative pollution content ratio in the post liberalization period suggest that the composition of our export bundle has shifted towards more pollution intensive goods and this composition effect has outweighed the technique effect. Between 1992 and 1996 there has been a slight decrease in the pollution content ratios. Both total and relative ratios have declined. The decline is more pronounced in case of the total pollution content ratio. This suggests that one factor behind the decline in the total pollution content ratio is the scale effect. The relative pollution content ratio has also decreased between 1992 and 1996.This suggests that between 1992 & 1996 the technique effect might have outweighed the composition effect. Both the total and the relative pollution content ratios have increased between 1996 and 2000.This increase in the pollution content ratios indicate that during this period, there has been increased specialization towards dirty industries, and this composition effect has outweighed the technique effect.

4. Concluding Observations

The major findings of our study can be summarized as follows:

i. Most of the commodity groups whose RCA indices have shown improvements in the post liberalization period are dirty.

ii. After liberalization, the pollution content of India's trade has increased significantly[6]:

---

[6] This kind of study conducted for four isolated years may sometimes give wrong idea about the actual trend. Hence policy conclusions cannot be drawn solely on the basis of this result. But the increase in pollution content and the magnitude of increase (143.75% increase in the total pollution content &71.26% increase in relative pollution content ratio between 1985 & 2000) that our results suggest, point towards the necessity of conducting further studies on the trade off between trade and environment in the context of India.

the total pollution content ratio has increased from 0.480 in 1985 to 1.17 in 2000.The relative pollution content ratio has increased from 0.835 in 1985 to 1.43 in 2000. Thus, from a net exporter of pollution India has become a net importer of pollution.

From the foregoing discussion it is apparent that in a globalized, liberalized set up, a less developed country can gain from trade only at the cost of specializing in environmentally intensive products. It is interesting to note that for a developed country like the US the pollution content ratio is always very close to one, though the magnitude of the ratio is falling over time (Robison 1988). Robison (1988) found that the US is a net exporter of pollution and the degree of this export is going up over time. For a developing country like India, the ratio is expected to be lower than unity indicating country status of net exporter of pollution. For India this ratio was less than one in the pre liberalization period but it has increased gradually since then and at present the ratio is greater than one. Thus from a net exporter of pollution India has become a net importer of pollution. In fact, the speed of increase for India in the post liberalization phase is a matter of serious concern. If this tendency cannot be reversed in near future the liberalization growth path may not be sustainable in the long run.

On the basis of our data set distinction could not be made between the specialization effects and migration effects. It would be interesting to study separately the effect of liberalization on pollution trading due to specialization in environmentally intensive products and due to migration of dirty industries to the South.

Antweiler, Copeland and Taylor (2001) suggested an integrated framework to consider more comprehensive hypotheses related to the pollution haven effect and factor endowment effect simultaneously. I intend to extend my framework in future to test Antweiler –Copeland –Taylor type comprehensive hypotheses.

Appendix

Data Matching

The first task in this data matching exercise is to find the export (Ei) & import (Mi) figures belonging to the different I-O categories. To do this we will have to match the different SITC categories to the I-O categories. This has been done following the CSO's sector description enclosed in the appendix III& IV of the CSO's transactions tables for 1983-84, 1989-90, 1993-94 and 1998-99. These tables record the different industries which are included under each I-O category. Each SITC category has been matched to the I-O sector, which includes the industry which manufactures the commodity/ commodities belonging to the SITC category. Take for example SITC categories 121 (Tobacco Unmanufactured Refuse) and 122 (Tobacco Manufactured). Products of these categories were exported by India in 1985. Both these categories have been included under the I-O commodity group 15 (Tobacco Products) in the

Abatement Cost Table for 1985. According to CSO's sector description I-O category 15 (Tobacco Products) includes tobacco stemming, redrying, grading etc and manufacture of bidi, cigars, cigarette etc. Detailed description regarding the matching of the SITC categories with the I-O categories is recorded in the abatement cost tables for the different years.

Generally each SITC category could be matched to a single I-O category. Problem arises with SITC 724 (textile, leather machinery) and SITC 728(Other Machi. for Special Industry). SITC 724 category can be included under both I-O 39(Industrial Machinery for Food and Textiles) and 40(Other machinery). It is also not clear from CSO's sector description whether 728 should be included in I-O 39 or 40. To remove these problems sectors 39 & 40 of CSO's input output transactions table has been aggregated, so that our sector 40 includes Machinery for Food and Textiles as well as other machinery. So, ultimately we are considering a 59×59 input output transactions matrix.[7]

The next step of the exercise would be to find out the abatement cost of the I-O sectors i. e., the C_js. The abatement cost data is available in NIC categories. The trade data is available in the SITC categories. The SITC categories have been matched to the NIC categories. Take for example SITC category 011(Meat Fresh Chilled). This has been matched to NIC category 200 (Slaughtering, Preparation &Preservation of meat). For most of the SITC categories this matching is possible but this scheme failed for the SITC categories belonging to I-O sectors 9&10[8]. For the SITC categories belonging to sectors 9&10 the abatement costs could not be accurately determined due to the non availability of the abatement cost data for the corresponding NIC categories. For example consider Crude Petroleum (333), included in I-O sector 9. Abatement cost data for the NIC category which corresponds exactly to this SITC category is not available. So, SITC category 333 has been matched to the NIC category, Mfg. Of refined petroleum Products (314), with abatement cost 0.0007 as 314 is the closest NIC category for which abatement cost data is available. The abatement cost of any I-O sector will obviously depend on the different commodities which are included in the I-O sector. Each I-O category contains SITC categories with different levels of abatement costs. So, to find out the abatement cost of any I-O category first we have to find out the abatement costs of all the SITC categories contained in the I-O category and then a weighted average of the abatement costs of the relevant categories have to be taken. In this exercise the shares of the different SITC categories in the total value of trade in each I-O category has been noted. Then to find the abatement cost of each I-O category the weighted average of the abatement cost of SITC categories included in the I-O category has been taken. The weights are the shares of the different SITC categories in the total value of trade of each I-O category. Though the abatement cost per SITC category is unchanged, as the constituent commodities and their quantitative importance are changing over the years, the I-O category wise abatement cost is also varying over time.

---

[7] It may be noted that the aggregation of sectors 39 & 40 is not expected to affect the results as the same scheme of aggregation has been followed for all the years under consideration.

[8] It may be noted that the abatement costs of the sectors 9 & 10 have remained constant over the period of our analysis. Constancy of the abatement cost figures indicates that the constituent SITC categories and their quantitative importance (i.e. the share of the constituent SITC categories in the total value of trade of the respective I-O sectors) have remained constant over the period of our analysis. So, the abatement cost figures assigned to these sectors have not affected the direction in which the pollution content ratios have moved. Of course, the magnitude of the ratios might have been affected. Thus, the direction of change would be preserved though the exact magnitude might change with more accurate data.

In the text, the methodology for finding the abatement cost of individual I-O categories has been illustrated with the abatement cost calculations for I-O category 13. For all other I-O sectors which include multiple SITC categories the same methodology has been used. Some I-O categories include a single SITC category or multiple SITC categories all of which has been matched to a single NIC category. For these I-O categories there is only one abatement cost that is relevant. For e.g., in 2000 the only SITC category included under I-O category 21 was SITC 821(Furniture and parts thereof). This SITC category has been matched to NIC Category 276(Manu. of Wooden Furniture and Fixtures) with abatement cost 0.00. I-O 21 has been assigned abatement cost 0.00 in the abatement cost table for 2000. In the abatement cost table for 2000 I-O 29 includes SITC categories 511(Hydrocarbons, nes and derivatives), 512(Alcohols, phenols etc and their derivatives), 513(Carboxylic acid and their derivatives), 514(Nitrogen-function compounds), 515(Organo-inorganic &heterocyclic compounds) and 516(Other organic chemicals). All these SITC categories have been matched to a single NIC category Organic and Inorganic Chemicals with abatement cost 0.0017. There are some I-O sectors whose products are not included in the export import list. For these I-O sectors the abatement cost of the NIC category, which corresponds to the I-O category in question, has been used. For example take I-O 33 (Cement). Products of I-O 33 were not exported or imported in 1985, 1992 1996 or 2000. This sector has been matched to the NIC category Mfg of cement Lime and plaster (324) with abatement cost 0.0007 to capture the linkage effect.

References

Antweiler W., Brian R. Copeland and M Scott Taylor (2001) : "Is Free Trade Good For The Environment?" *The American Economic Review, Vol 91, No 4.*

Cole Mathew A., (2000) : Trade liberalisation, economic growth and the environment, *Edward Elgar Publishing, Inc.*

Cole Mathew A., (2004) : US environmental load displacement : examining consumption, regulations and the role ofNAFTA. *Ecological Economics Vol 48.*

Copeland Brian R., and M Scott Taylor, (1995) : Trade and Transboundary Pollution. *The American Economic Review, Vol. 85, No 4.*

Copeland Brian R., and M. Scott Taylor, (1994) : North- South Trade and the Environment. *The Quarterly Journal of Economics, August.*

Dean Judith M., (1992) : "Trade and the Environment: A Survey of the Literature. in: P Low ed., International Trade and the Environment. *World Bank Discussion Paper No. 159 (World Bank, Washington D.C., 1992)*

Dorfman, R., P.A.Samuelson, and R. Solow (1958): "Linear Programming and Economic Analysis"

Low P., and A. Yeats, (1992) : Do "Dirty" Industries Migrate? in: P Low ed., International Trade and the Environment. *World Bank Discussion Paper No. 159 (World Bank, Washington D.C., 1992)*

Lucas, R.E.B., D. Wheeler and H. Hettige, (1992) : Economic Development, Environmental Regulation and the International Migration of Toxic Industrial Pollution. in: P Low ed., International Trade and the Environment. *World Bank Discussion Paper No. 159 (World Bank, Washington D.C., 1992)*

Mani M., and D. Wheeler, (1999) : In Search of Pollution Havens? Dirty Industry in the World Economy, 1960-1995. in: Trade, Global Policy and the Environment. *World Bank Discussion Paper No 402 (World Bank, Washington D. C., 1999)*

Muradian,R.,O'Connor,M.,Martinez-Alier,J., (2002) : Embodied pollution in trade:estimating the 'environmental load displacement' of industrialized countries.*Ecological Economics Vol 41.*

Robison David H., (1988) : Industrial Pollution Abatement: The Impact on Balance of Trade. *Canadian Journal of Economics. (February 1988)*

Tobey James A., (1990) : The Effects of Domestic Environmental Policies on Patterns of World Trade: An Empirical Test. *Kyklos, Vol. 43.*

Arthaniti 5 (1-2), 96-113 / 2006

Monetary Models of Exchange Rates and The Random Walk

Gianluca Laganà* and Pasquale M Sgro**

Abstract

Three alternative monetary models of exchange rate are tested using data on the Italian lira - US dollar exchange rate. It is shown that up to the early 1990s these economic models perform better than the random walk model in out-of-sample forecasts.

Keywords : monetary models of exchange rate, random walk with drift, out-of-sample forecasts, efficient market hypothesis.

JEL Classification : F310, F470.

1. Introduction

One of the main themes of empirical international economics is the role that macroeconomic models play in explaining and predicting the short-run and/or the long-run behaviour of the nominal exchange rate.

Historically, the main empirical results in the literature demonstrated that out-of-sample asset market models do not predict the exchange rate significantly better than a naïve random walk (RW) model (Meese and Rogoff (1983a), (1988)[1,2]. In particular, Meese and Rogoff (1983a) found that a simple RW model outperforms all structural models in out-of-sample forecasting of the exchange rate for four Western countries[3].

The response to these negative results was mixed. Some used the monetary models of exchange rates (MMER) with time-series and panel-data (Mark (1995),

---

* Facoltà di Giurisprudenza, Università Roma Tre, Via Ostiense 161, 00154 – Roma – Italy; e-mail: lagana@giur.uniroma3.it.

** Deakin Business School, Deakin University, 336 Glenferrie Road, Malvern, Victoria 3144, Australia: email sgro@deakin.edu.au

[1] The main reason for the relevance of macroeconomic fundamentals for forecasting exchange rate movements is given by the asset market approach of exchange rate. Under this approach, demand for currencies is determined by purchases and sales of goods as well as assets.

[2] Fitting exchange rates to contemporary observable variables, in sample, is one thing. Forecasting out-of-sample is quite another. A stronger test of the MMER's validity is to determine how well they perform out-of-sample compared to an alternative, such as the naive RW with drift.

[3] Since the work of Meese and Rogoff (1983a) the focus of the attention of the literature on the determination of the exchange rate has changed from in-sample estimation to out-of-sample forecasting. These results are also confirmed by Cushman(2000) and Meese (2001)

Blomberg and Hess (1997), Mark and Sul, 2001)[4]; b) others like Gandolfo et al., (1990), (1992)[5] used large-scale macroeconomic models. A third avenue of approach was the use of non-linear models, Balke and Forby (1997), Kilian and Taylor (2003).[6]

Recently more sophisticated technical tools have been used to support the argument that economic models forecast exchange rates better than random walk models. Rossi (2006), (2005) for example, shows that allowing for parameter instability in models of nominal exchange rate determination improves the forecast relative to the random walk.[7] Guo and Savickas (2005) on the other hand use the notion of "aggregate idiosyncratic volatility" to forecast the future US dollar rates against various currencies. These results also support the monetary models.

This paper contributes and builds on this debate by using a simple monetary model in which there are market imperfections, such as capital controls, and where we show that a number of monetary models can outperform the random walk models. The monetary models that are used are (1) the flexible-price monetary model due to Frenkel (1976) and Bilson (1978), (2) the Dornbusch-Frankel monetary model with sticky prices and (3) the uncovered interest rate parity condition. We use the Italian lira – US$ exchange rate to compare the forecasting ability of these different exchange rate models with a benchmark random walk model using out-of-sample predictions. In other words, the ability of each competing model to generate out-of-sample forecasts at one- to twelve-month horizons for the Italian lira – US dollar exchange rate is assessed. The parameters of each model are estimated for every forecast period by the rolling regressions method using the most up-to-date information available at the time of a given forecast.

Our main results indicate that first pre- 1992, the MMER is a good predictor of the Italian Lira – US $ exchange rate in out-of-sample forecasting. In particular, in out-of-sample forecasting, MMERs significantly outperform the benchmark RW model up to a twelve-months horizon. These findings can be interpreted in light of the presence of controls on capital movements in Italy until the early 90s[8] and second post- 1992, the MMER ceases to be a good predictor of the exchange rate.

---

[4] Mark (1995) and Chinn and Meese (1995) proved the usefulness of monetary models in predicting the nominal exchange rate in the long-term; Blomberg and Hess (1997) demonstrated that monetary models augmented with political variables out-perform the random walk model in the short-term forecasting ability of the nominal exchange rate; Mark and Sul (2001) reported positive short-term results for a panel of countries.

[5] Gandolfo, Padoan and Paladino (1990) abandon the idea of using traditional single equation structural models of exchange rate, favouring instead the use of economy-wide macro-econometric models, and found positive results for the Italian lira versus the US dollar nominal exchange rate.

[6] Since the literature on these studies is so vast it is not our intention to provide here a summary of it (for further details refer to Satrno and Taylor (2002) and Franses and van Dijk (2000).

[7] Clark and McCraken (2002) have also analyzed model selection and forecasting with parameter instability.

This, in part, contradicts the MMER framework, which argue that, in perfectly competitive markets , economic models should perform better than RW. The Meese-Rogoff (1983a) negative results hold for Italy (just after the departure of the Lira from the EMS in 1992).

The robustness of these results is assessed by introducing forward-looking variables (expected inflation and survey data on production expectation) into these models and by changing the sample period. Results appear to be robust to these changes[9].

This paper is organized as follows: section 2 describes the methodology used to compare different models; section 3 describes the competing exchange rate models that need to be tested out-of-sample; the sub-section 3.1 derives the benchmark RW with drift; section 4 of the paper provides a preliminary data analysis; section 5 contains the out-of-sample results; the issue of robustness is addressed in section 6 and the conclusion follows in section 7.

2. Methodology

This section presents the methodology used to assess the forecasting ability of MMERs against that of a benchmark model.

The out-of-sample forecasts are obtained by using recursive estimation methodology. The data is partitioned in such a way that the proposed models are first estimated using data from January 1974 to December 1984. The estimated coefficients are then used to construct the one-step-ahead forecast of the bilateral exchange rate for the country considered in this paper. The regressions are then re-estimated with one more observation, using data from 1974:1 to 1985:1, to obtain the forecast for 1985:1. The process is then continued in this recursive manner until the last forecast (1988:12) is constructed. The process is then repeated for the 2-, 3-, 4-, 5-, 6- and 12-months-ahead prediction error. This recursive methodology produces a vector of 48 forecast errors for the 1-month-ahead prediction error, another of 47 for the 2-months-ahead prediction error, and so on until the 12 months-ahead prediction error is constructed. Each vector is then used to calculate the Root Mean Squared Error

---

[8] The controls on the capital movements in Italy and their implications are discussed extensively in Berardi, Dalmazzo e Marini (1996); Pozzolo (1999) and Ciocca (2000).

[9] Gandolfo et al. (1990) also found, using a different dataset and a multiple-equation model "MARK V", that the benchmark RW model can be out-performed in out-of-sample forecasting by a wide-macro-econometric model for the Italian lira – US dollar exchange rate. Gandolfo et al. (1990), however, stressed that single equation monetary models were not useful in predicting exchange rate movements.

(henceforth RMSE), which is a standard measure for assessing competing models in out-of-sample forecasts:

$$\mathrm{RMSE} = \left\{ \sum_{s=0}^{Nk-1} \left[F(t+s+k) - A(t+s+k) \right]^2 2/Nk \right\} \tag{1}$$

where k = 1, 2, 3, 4, 5, 6, 12 denotes the forecast step, N_k the total number of forecasts in the projection period for which the actual value A(t) is known and F(t) the forecast value. Forecasting begins in period t (see Meese and Rogoff 1983a).

3. The model

The aim of this paper is to report new empirical results on the Italian lira - US dollar exchange showing that a simple univariate-linear model, which includes macroeconomic fundamentals, is able to outperform a benchmark RW with drift in predicting the exchange rate. For this purpose, two versions of the monetary model were selected. The first is the flexible-price monetary model, originally introduced by Frenkel (1976) and Bilson (1978). The second is the Dornbusch-Frankel monetary model with sticky prices. A third model, based on the uncovered interest parity condition, is also considered. All models are assessed against a benchmark RW with drift. Consider the following specification:

$$s_t = a_0 + a_1(m-m^*)_t + a_2(y-y^*)_t + a_3(i-i^*)_t + a_4(\pi-\pi^*)_t + u_t \tag{2}$$

All variables are in logs and to conform to the empirical literature on exchange rates the dependent variable is considered in log first difference. Jensen's inequality problem is avoided by comparing predictors on the basis of their ability to predict the logarithm of the exchange rate. In other words, because of Jensen's inequality, the best predictor of the level of the Italian lira – US dollar exchange rate might not be the best predictor of the US dollar – Italian lira exchange rate. s_t, is the difference in the spot exchange rate defined as Italian lira per one US dollar, y_t, is the output level, i_t, is the short- term interest rate, π_t is the consumer price index and u_t is a random error. An asterisk next to a variable indicates foreign values. Equation (2) makes the assumption that domestic and foreign variables affect the exchange rate with coefficients of equal but of opposite sign.

The flexible-price monetary model of exchange rates (henceforth Model 2) is readily obtained from equation (2) by restricting a_1 to be unity, $a_2<0$, $a_3>0$ and $a_4=0$. The sticky-price monetary model (Model 1) is obtained by imposing the following restrictions: $a_1=1$, $a_2<0$, $a_3<0$ and $a_4>0$. The main difference between the flexible- and sticky-price monetary models is that the DF model allows for short-term deviation from purchasing power parity in the goods market. Our proposed model, based on the uncovered interest parity condition (henceforth Model 3), uses the following restrictions: $a_1=a_2=a_4=0$ and $a_3\neq 0$. Lastly, as a benchmark model, a RW model with drift (this is explained below), is based on the restrictions $a_1=a_2=a_3=a_4=0$ is used (Model 4).

The reason for incorporating the interest rate differential in the model (Model 3) is based on the idea that the adjustment of interest rates influences the movements in the exchange rates. So, for instance, by raising the domestic interest rate, the foreign exchange value of the domestic currency unit could be strengthened. On the other hand, by reducing the domestic interest rate the domestic currency could depreciate. From the above equation it is possible to see that an increase in the domestic money supply, relative to the foreign money supply, depreciates the domestic currency. An increase in the domestic output increases the value of the domestic currency (through PPP). All variables are dated as if they were known ex-ante.

3.1 Derivation of the Random Walk with drift

Numerous empirical studies of international financial markets have focused on market efficiency. In an efficient speculative market, prices should fully reflect information available to market participants and it should be impossible for a trader to earn excess returns on speculation. Foreign exchange market participants possess two characteristics: (i) rational expectation; (ii) risk neutrality. The efficient market hypothesis in the presence of risk neutrality implies that the gain from holding one currency rather than another must be compensated for by the opportunity cost of holding investments in this currency rather than the other. Black (1971) defined a perfect market for a stock as one in which both people endowed with publicly available information and those with private information are unable to make profits from speculation (because prices adjust very quickly as the information becomes available, therefore, prices move randomly). To illustrate this more clearly, consider the concept of a Fair-Game. A Fair-Game is a game which is neither in your favour nor in that of your opponent. This is the essence of a martingale, a stochastic process $[s_t]$ which satisfies the following condition:

$$E[s_{t+1}-s_t \mid s_t, s_{t-1}]=0 \tag{3}$$

Suppose that the variable s_t can be written as the rational expectation of some "fundamental value" V*, conditional on all available information at time t "I_t". Formally:

$$s_t = E[V^*|I_t] = E_tV^* \tag{4}$$

It follows that:

$$s_{t+1} = E[V^*|I_{t+1}] = E_{t+1}V^* \tag{5}$$

Since $I_t \subseteq I_{t+1}$, thus:

$$E_t[E_{t+1}[V^*]] = E_t[V^*] \tag{6}$$

All the information in I_t is also in I_{t+1}, but I_{t+1} is superior because it contains extra information. By the Law of Iterated Expectation:

$$E[s_{t+1}-s_t] = E_t[[E_{t+1}[V^*] - E_t[V^*]]=0 \tag{7}$$

In other words, realised price changes are unforecastable given the information in the set I_t. If s_t is the spot price of one country's money in terms of another at time t, then the martingale hypothesis states that the change in the spot price exchange rate is zero when conditioned on the exchange rate's entire history. In a forecasting context, the martingale hypothesis implies that the best (here best = minimal mean-squared error) forecast of tomorrow's exchange rate is simply today's exchange rate. Until recently the martingale hypothesis was considered to be a necessary condition for an efficient asset market. The more efficient the market, the more random the sequence of price changes generated by the market. Unfortunately the martingale hypotheses do not account for "risk" in any way, which is indeed one of the most important concepts in modern financial economics (trade-off risk-return). Specifically, if the change in the spot exchange rate is positive, it may be the reward necessary to attract investors to hold that currency and its associated risk. Thus implying that the martingale hypothesis is neither a necessary nor a sufficient condition for rationally determined asset prices [see Lucas (1978)].

To overcome the aforementioned problem, tests on the efficient market hypothesis are carried out on a stronger version of the martingale model; one which assumes identically and independently distributed increments. This model, known also as "random walk with drift", is given by the following expression: $s_t = c + s_{t-1} + e_t$ where c is the expected change in the exchange rate or drift.

4. Description of data

This section examines the sample moments of the variables considered in the paper. The first and the second moment of all variables used in the paper are reported. The raw data is monthly and is taken from DataStream. The full sample covers the period from January 1974 to November 1998. Variables are selected according to the asset market approach to exchange rates. The spot exchange rate, presented in Italian lira per one US dollar, is used as the exchange rate, the index of industrial production as a measure of output, the monetary aggregate m_1 as a measure of money supply, the money market interest rate as a measure of the short-term nominal interest rate and the consumer price index is used to calculate a measure of inflation. In this paper Italy represents the domestic country while the US represents the foreign country. Foreign variables are denoted with an asterisk. Table 1 provides the mean and standard deviation of all variables considered in this paper. Each column reports the variable name, the mean and the standard deviation over the full sample.

TABLE 1

Sample moments: 1974(1)-(11)1998

	s	y	π	m	i	y^*	π^*	m^*	i^*
mean	1285	85	8.1	262	12.9	80	4.6	676	7.7
variance	367	11	4.8	124	3.9	15	2.7	304	3.3

Notes: The variables in the table are defined as follows: s is the Italian lira – US dollar nominal exchange rate; y is output; π is the rate of inflation; m is the money supply and i is the short-term interest rate.

The main empirical regularity of this exchange rate is as follows: First, the Italian interest rates and inflation rates were on average almost twice those of the US rates. Second, the volatility of the Italian interest rates and inflation rates was greater than that of the foreign rates.

5. Results: Out-of-sample forecasts

The empirical analysis in this section focuses on the out-of sample estimation. We commence the preliminary data analysis by first looking at the cross-correlations between the series and then by choosing their optimal lag length for the estimation according to the Akaike information criteria. Then Model 1, Model 2 and Model 3 are assessed in an out-of-sample estimation for the Italian lira – US dollar bilateral exchange rate against a RW with drift. The main result of this section is that the MMER has a lower RMSE than the RW up to twelve-months ahead. This result is robust across different models and for different sample periods.

The out-of-sample forecast results are displayed in Table 2 which is organized in the following way: the forecasting horizon, which goes from 1 to 12 months, is contained in the first column of the table, the second column reports the RMSE results of the Dornbusch-Frankel (DF) model, the third column reports the RMSE results of the Frenkel-Bilson (FB) model, the fourth column reports the RMSE results of the model based on the interest rate differential and the fifth column reports the RMSE results of the benchmark RW model.

The issue of data mining is then addressed by repeating the experiment over a different sample period (1986:7-1990:6) to account for the fact that the above results could be dependent upon a particular sample period used in the estimation. Results are shown in the bottom quadrant of Table 2. The out-of-sample results for the bilateral exchange rates confirm our hypothesis, for the Italian lira – US dollar nominal exchange rate, that the simple MMERs have a lower RMSE than the benchmark RW. As can be seen from Table 2, during the period 1984:12 – 1988:12 all models outperform the RW significantly for up to a year. Repeating the experiment over the period 1986:6 – 1990:6 it is found, however, that Models (1)-(3) lose some of their advantage over the RW in predicting the exchange rate movements.

The out-of-sample forecasts based on the RMSE results confirm the previous finding that the MMERs outperform the RW. These results are consistent and robust across different sample periods.

TABLE 2

RMSE from out-of-sample-forecast

Nominal exchange rate Lira /$ - Sample: 1984(12)-(12)1988

	Model 1	Model 2	Model 3	R W
Horizon				
1	0.0251*	0 0250*	0 0265**	0.0297
2	0 0286***	0 0281***	0 0302*	0 0311
3	0 0293***	0 0287***	0.0307	0 0311
4	0.0298***	0.0290***	0.0311***	0.0313
5	0.0289*	0 0283***	0.0299**	0.0302
6	0 0293**	0.0286***	0.0303**	0.0305
12	0.0283**	0.0275**	0.0290**	0.0292

Nominal exchange rate Lira /$ - Sample: 1986(6)-(12)1990

	Model 1	Model 2	Model 3	R W
Horizon				
1	0 0224*	0.0226*	0 0232	0.0255
2	0 0242*	0.0244*	0.0255	0 0256
3	0.0238	0.0241	0 0248**	0.0249
4	0.0241	0 0244**	0.0251	0.0251
5	0.0242	0.0246	0.0252	0 0252
6	0 0245	0.0248	0.0255	0.0255
12	0.0247	0.0251	0.0255	0 0253

Notes:

Model 1, in addition to the dependent variable, includes a constant and lags of the differentials between domestic and foreign short-term interest rate, domestic and foreign money supply, domestic and foreign output and domestic and foreign inflation. Model 2 is simply defined as Model 1 less the inflation differential. Model 3 in addition to the dependent variable, includes a constant and lags of the differential between domestic and foreign short-term interest rates. The time horizon is reported in the first column, columns 2, 3, and 4 report the RMSE results of Models (1) – (3) and compare their significance level with the RMSE of the benchmark Random Walk Model. The asterisk beside the RMSE indicates the significance level of the Sign test, which is applied to the vector of each forecast error according to the formula S= S - 0.5T/sqrt(0.25T). S (which is asymptotically distributed with a standard normal N(0,1), and represents the number of positive forecast errors) and T the sample size (T=48 for the 1-step-ahead forecast and T=37 for the 12-step-ahead) [see Diebold and Mariano (1995)]. As pointed out by Diebold and Mariano (1995), the use of the F-test to compare the forecast errors series would be of little use here because the conditions required to obtain its distribution are too restrictive.

*different from random walk at a 10% significance level.

**different from random walk at a 5% significance level.

***different from random walk at a 1% significance level.

6. Robustness

The issue of robustness is explored along a number of dimensions. First, the Italian and US measures of inflation are replaced with those of expected inflation. Second, to provide a more stringent robustness test, the sample period for the out-of-sample estimation is changed and a measure of expected output from survey data is used instead of the Italian and US measures of output used earlier.

6.1. Expected Inflation

Table 3 reports the out-of-sample results once the measure of expected inflation is included and replaces inflation. Expected inflation introduces a forward-looking element into the analysis.

TABLE 3

RMSE from out-of-sample-forecast

Nominal exchange rate Lira /$ - Sample: 1984(12)-(12)1988

	Model 1	Model 2	R W
Horizon			
1	0.0264***	0.0276***	0.0297
2	0.0323	0.0322	0.0311
3	0.0335	0.0327	0.0311
4	0.0342	0.0331	0.0313
5	0.0324	0.0309	0.0302
6	0.0326	0.0311	0.0305
12	0.0304	0.0291	0.0292

Nominal exchange rate Lira /$ - Sample: 1986(6)-(12)1990

	Model 1	Model 2	Model 3	R W
Horizon				
1	0.0226***	0.0231*	0.0255	
2	0.0241***	0.0248***	0.0255	
3	0.0236	0.0239	0.0249	
4	0.0239***	0.0241***	0.0251	
5	0.0242***	0.0244***	0.0252	
6	0.0244***	0.0246***	0.0256	
12	0.0250	0.0248	0.0253	

Notes:

To introduce a forward-looking element into the analysis, in Model 1 the Italian – US expected inflation differential replaces the inflation differential. Model 1, in addition to the dependent variable, also includes a constant and lags of the differentials between domestic and foreign short-term interest rates, domestic and foreign money supply, domestic and foreign output and domestic and foreign expected inflation. Model 2, in addition to the dependent variable, includes a constant and lags of the differential between domestic and foreign expected inflation. The time horizon is reported in the first column, columns 2, 3, and 4 report the RMSE results of Model (1) – (2) and compare their significance level with the RMSE of the benchmark Random Walk Model. The asterisk beside the RMSE indicates the significance level of the Sign test, which is applied to the vector of each forecast error according to the formula S= S - 0.5T/sqrt(0.25T). S(which is asymptotically distributed with a standard normal N(0,1), represents the number of positive forecast errors) and T the sample size (T=48 for the 1-step-ahead forecast and T=37 for the 12-step-ahead) [see Diebold and Mariano (1995)]. As pointed out by Diebold and Mariano (1995), the use of the F-test to compare the forecast errors series would be of little use here because the conditions required to obtain its distribution are too restrictive.

*Different from random walk at a 10% significance level.

**Different from random walk at a 5% significance level.

***different from random walk at a 1% significance level.

As before, the out-sample-forecasting exercise is carried out on two different sample periods (1984:12-1988:12; 1986:7-1990:6). The results from the first sample period (1984:12-1988:12) show that Models 1 and 2 both significantly outperform the RW with drift at one-month-ahead. Those from the second (1986:7-1990:6) report a striking superiority of fundamentals based models over the RW (Model 1 and Model 2 outperform the RW up to six-months-ahead)[10].

6.2 Expected Output

A more stringent robustness test is: 1) to change the sample period considered for the –out-of-sample forecasting exercise and 2) to introduce a forward-looking element into the analysis in the place of the actual output. This forward-looking element is represented by "production expectation for Italy" and is taken from the European Commission Business Survey. The European Community business survey is carried out in Italy by the ISAE (Istituto Studi e Analisi Economica, previously ISCO) national institution on the basis of the Commission's harmonized questionnaires. The EU survey is not designed to yield exact quantitative forecasts but takes the form of a trend survey in which respondents may choose from 3 categories of responses. During the last ten days of each month around 20,000 firms (all in the EU) are contacted by national institutions (such as the ISCO) and asked to respond to a set of questions, including "Production expectations for the month ahead: up, unchanged, down?". The majority of replies are received by the first week of the month to which the questionnaires refer[11]. Results are then published during the first week of the following month. The survey data may thus be interpreted as an assessment of the movement of the respective national production. Despite the existence of other income variables such as GNP, GDP and the Index of Industrial Production, the advantage of this study is that the time required for data collection and release is very short. In the foregoing robustness analysis the variable representing production expectation for the month ahead is converted from a qualitative to a quantitative measure in the following way:

---

[10] MacDonald and Marsh, (1997) present evidence to support the proposition that fundamentals based models have clear ability to forecast at horizons as short as two months ahead; Mark and Soul (2001) find that monetary fundamentals have a small but statistically significant level of predictability for the exchange rate.

- measure 1 = $(\text{Up} + \text{Unchanged}/2)_{t-1}$;
- measure 2 = $(\text{Up} - \text{Down})_{t-1}$;
- measure 3 = $(\text{Up} + \text{Unchanged}/2)_{t-2}$

The results from the out-sample-forecasting exercise are reported in Table 4. These results are based on the following sample period (1988:7-1992:6). The first column indicates the forecasting horizon, the following two columns report the RMSE results using the first measure of output expectation (and are contrasted with those of the benchmark RW with drift) and the remaining columns show the RMSE results obtained using the other two measures of production expectation as explained above.

The results show that Models 1 and 2 both significantly outperform the RW with drift at two-months-ahead. Furthermore, these results are not dependent on the measure of expected output used.

The robustness of the results confirms our previous finding that for the Italian lira – US dollar nominal exchange rate the MMERs outperform the RW.

---

[11] The only exceptions are Italy and Germany, where the questionnaires are issued around the 15th – 20th of the month preceding that to which the result are attributed. In both cases, however, most of the completed questionnaires are received during the first 10 days of the month to which the results are attributed.

TABLE 4

RMSE from out-of-sample-forecast

Nominal exchange rate Lira /$ - Sample: 1988(7)-(6)1992

	Model 1	Model 2	R W	Model 1	Model 2	R W	Model 1	Model 2	R W
	(using measure 1)			(using measure 2)			(using measure 3)		
Horizon									
1	0.02517***	0.02474***	0.02721	0.02529***	0.02484***	0.02721	0.02528***	0.02492***	0.02721
2	0.02658*	0.02552***	0.02685	0.02667*	0.02563***	0.02685	0.02672**	0.02950	0.02685
3	0.02639	0.02590	0.02701	0.02651	0.02604*	0.02701	0.02656	0.02635	0.02701
4	0.02698	0.02661	0.02728	0.02708	0.02675	0.02728	0.02712	0.02702	0.02728
5	0.02691	0.02671	0.0272	0.02704	0.02688	0.0272	0.02708	0.02715	0.0272
6	0.02632	0.02627	0.02652	0.02648	0.02648	0.02652	0.02655	0.02676	0.02652
12	0.02701	0.02708	0.02762	0.02712	0.02727	0.02762	0.02710	0.02741	0.02762

Notes:

To introduce a forward-looking element into the analysis, in Model 1 the Italian – US expected output differential replaces the output differential. The three measures used in the table are the following: measure 1 = $(\text{Up} + \text{Unchanged}/2)_{t-1}$; measure 2 = $(\text{Up} - \text{Down})_{t-1}$; measure 3 = $(\text{Up} + \text{Unchanged}/2)_{t-2}$ Model 1, in addition to the dependent variable, includes a constant and lags of the differentials between domestic and foreign short term interest rate, domestic and foreign money supply, domestic and foreign inflation and domestic and foreign expected output. Model 2, in addition to the dependent variable, includes a constant and lags of the differential between domestic and foreign expected output. The time horizon is reported in the first column, columns 2, 3, and 4 report the RMSE results of Model (1) – (2) and compare their significance level with the RMSE of the benchmark Random Walk Model. The asterisk beside the RMSE indicates the significance level of the Sign test, which is applied to the vector of each forecast error according to the formula S= S - 0.5T/sqrt(0.25T). S(which is asymptotically distributed with a standard normal N(0,1), represents the number of positive forecast errors) and T the sample size (T=48 for the 1-step-ahead forecast and T=37 for the 12-step-ahead) [see Diebold and Mariano (1995)]. As pointed out by Diebold and Mariano (1995), the use of the F-test to compare the forecast errors series would be of little use here because the conditions required to obtain its distribution are too restrictive.

*different from random walk at a 10% significance level.

**different from random walk at a 5% significance level.

***different from random walk at a 1% significance level.

6.3 A further robustness check

The aim of this section is to assess whether our hypothesis is correct: that monetary models, with both flexible and sticky prices, outperform the random walk model in out-of-sample forecasts before, but not after, 1992 – when the Italian lira departed from the EMS and most of the legal restrictions to capital movements were eased. In order to do this, the experiments carried out in section 6.2 were repeated using a sample of data that went beyond 1992.

Results from the out-sample-forecasting exercise are reported in Table 5. The Table can be read as before (section 6.2), the only difference being that the results are now carried out on the following sample period (1992:1-1994:6). The results, in sharp contrast with all previous findings, show that both Model 1 and Model 2 no longer outperform the benchmark RW, thus implying that fundamentals are unable to predict exchange rate movements.

TABLE 5

RMSE from out-of-sample-forecast

Nominal exchange rate Lira /$ - Sample: 1992(1)-(6)1994

	Model 1	Model 2	R W	Model 1	Model 2	R W	Model 1	Model 2	R W
	(using measure 1)			(using measure 2)			(using measure 3)		
Horizon									
1	0.02929*	0.03035	0.03333	0.02939	0.03037	0.03333	0.02932*	0.03030	0.03333
2	0.03373	0.03952	0.03366	0.03374	0.03939	0.03366	0.03375	0.03927	0.03366
3	0.03393	0.04199	0.03385	0.03392	0.04186	0.03385	0.03393	0.04168	0.03385
4	0.03244	0.04169	0.03244	0.03239	0.04155	0.03244	0.03241	0.04132	0.03244
5	0.03187	0.04150	0.03204	0.03181	0.04136	0.03204	0.03183	0.04109	0.03204
6	0.03212	0.04158	0.03236	0.03206	0.04143	0.03236	0.03208	0.04115	0.03236
12	0.03275	0.04320	0.03319	0.03265	0.04302	0.03319	0.03267	0.04268	0.03319

Notes:

To introduce a forward-looking element in the analysis, in Model 1 the Italian – US expected output differential replaces the output differential. The three measures used in the table are the following: measure 1 = $(\text{Up} + \text{Unchanged}/2)_{t-1}$; measure 2 = $(\text{Up} - \text{Down})_{t-1}$; measure 3 = $(\text{Up} + \text{Unchanged}/2)_{t-2}$

Model 1, in addition to the dependent variable, includes a constant and lags of the differentials between domestic and foreign short term interest rate, domestic and foreign money supply, domestic and foreign inflation and domestic and foreign expected output. Model 2, in addition to the dependent variable, includes a constant and lags of the differential between domestic and foreign expected output. The time horizon is reported in the first column, columns 2, 3, and 4 report the RMSE results of Models (1) – (2) and compare their significance level with the RMSE of the benchmark Random Walk Model. The asterisk beside the RMSE indicates the significance level of the Sign test, which is applied to the vector of each forecast error according to the formula S= S - 0.5T/sqrt(0.25T). S(which is asymptotically distributed with a standard normal N(0,1), represents the number of positive forecast errors) and T the sample size (T=48 for the 1-step-ahead forecast and T=37 for the 12-step-ahead) [see Diebold and Mariano (1995)]. As pointed out by Diebold and Mariano (1995), the use of the F-test to compare the forecast errors series would be of little use here because the conditions required to obtain its distribution are too restrictive.

*different from random walk at a 10% significance level.

**different from random walk at a 5% significance level.

***different from random walk at a 1% significance level.

In particular we argue that, after the early 90s, fundamentals-based macroeconomic models are unable to beat the naïve RW in predicting the Italian lira – US dollar exchange rate using out-of-sample forecasting. We explain these findings in light of the departure of the Italian lira from the European Monetary System (EMS) and the subsequent easing of controls on capital movements in Italy.

7. Conclusion

The paper has presented alternative simple monetary models of exchange rate determination which out-performed the benchmark random walk model in out-of-sample forecasts. Three alternative versions of the monetary model are presented and tested. The first is a simplified version of the sticky-price monetary model. The second is a simplified version of the flexible-price monetary model. The third is based on the interest rate differential. The performance of these models has been assessed against a RW with drift. Results indicate that the MMERs outperform the RW in predicting the Italian lira – US dollar rate of exchange at one-to-twelve months' horizons until the early 90s. The superiority of the MMER with respect to the RW model vanishes toward the end of 1992 (after the departure of the Italian Lira from the European Monetary System - EMS).

Our paper therefore adds to that body of literature where the economic model outperforms the RW and, in our case, these findings are most likely dependent on the presence of market imperfections such as capital control and high inflation in the Italian economy during the period. This feature is important since the presence of market imperfections for most economies is the norm rather than the exception. These findings appear to be robust across models of different specifications and over different sample periods. Finally simple models are still powerful tools in helping us understand observed market behaviour.

Reference

Berardi, A., Dalmazzo, A. and G. Marini, (1996) "Regolamentazione di flussi di capitale e volatilità del tasso di cambio. Risultati preliminari sull'esperienza italiana" *Rivista di Politica Economica,* Anno LXXXVI, Serie III pp. 119-137.

Black, F. (1971) "Toward a Fully Automated Stock Exchange", *Financial Analysts Journal*, July-August, 29-44.

Blomberg, S. B. and G. Hess, "Politics and Exchange Rate Forecasts", *Journal of International Economics*, 43, 189-205, 1998.

Campbell, J., Low, A. and C. Mackinlay, The Econometrics of Financial Markets, Princeton University Press, 1997.

Ciocca, P., "La nuova finanza in Italia. Una difficile metamorfosi (1980-2000)", Bollati Boringhieri, Torino 2000.

Clark, T.E. and M.W. McCracken (2002) "The Power of Tests of Prediction .... in the Presence of Structural Breaks" Journal of Econometrics.

Diebold, F. X. and R. Mariano, (1995) "Comparing Predictive Accuracy", *Journal of Business and Economic Statistics*, 13, 253-63.

Engel, C. and K. West (2004) "Exchange rate and Fundamentals" Journal of Political Economy.

Fama, E. (1970) "Efficient Capital Markets: A Review of Theory and Empirical Work", *Journal of Finance*, 25, 383-417.

Frankel, Jeffrey A., and A. K. Rose, "Empirical Research on Nominal Exchange Rates" in Handbook of International Economics, vol. III, Edited by G. Grossman and K. Rogoff, Elsevier Science B. V.,1995.

Frankel, Jeffrey A., On exchange rates Cambridge, MA: MITT Press, 1993.

Froot, K. A., "Using Survey Data to Test Standard Propositions Regarding exchange Rate Expectations", published in Frankel, MIT press, 1993.

Gandolfo, G., "International Economics II: International Finance and Open-Economy Macro-Economics", Berlin: Springer-Verlag, 1998.

Gandolfo, G., Padoan C.P. and G. Paladino (1990) "Structural Models Vs Random Walk: The Case of the Lira/$ Exchange Rate", *Eastern Economic Journal*, Vol. XVI, 101-123.

Gandolfo, G., Padoan C.P. and De Arcangelis, G. "The Theory of Exchange Rate Determination, and Exchange Rate Forecasting", CIDEI working paper N.1, June 1991.

Guo, H. and R. Savickas (2005), "Foreigh Exchange Rates Don't Follow a Random Walk" : The Federal Reserve Bank of St. Louis Working Paper 2005-025.

Kilian, Lutz (1999) "Exchange Rates and Monetary Fundamentals: What Do We Learn from Long-Horizon Regressions?" *Journal of Applied Econometrics*, 14(5), pp. 491-510.

Levich, R. (1985) "Empirical Studies of Exchange Rates: Price behavior, Rate Determination and Market Efficiency", Handbook of International Economics, vol. II, edited by R. Jones and P. Kenen, Elsevier Science B. V.

MacDonald, R. (1998a), Floating Exchange Rates: Theories and Evidence, London: Unwin-Hyman.

Mark, N.C., "Exchange Rates and Fundamentals; Evidence on Long-Horizon Predictability", *American Economic Review*, 85, 201-218, 1995.

Mark N.C. and D. Sul (2001) "Nominal Exchange Rates and Monetary Fundamentals: Evidence from a Small Post-Bretton Woods Panel." *Journal of International Economics*, February, 53(1), pp. 29-52,

Meese, R. (1990) "Currency Fluctuations in the Post-Bretton Woods Era", *Journal of Economic Perspectives*, 4, 117-134.

Meese, R. and K. Rogoff (1983b), "The Out-of-.Sample Failure of Empirical Exchange Rate Models: Sampling Error or Misspecification?" in J.A. Frenkel (ed.) Exchange Rates and International Macroeconomics, Chicago: Chicago University Press and National Bureau of Economic Research pp. 67-105.

Meese, R. and K., Rogoff (1983a) "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample", *Journal of International Economics*, 14(2), pp. 3-24.

Obstfeld, M. and K. Rogoff, "The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?", NBER, Macroeconomics Annual, 2000.

Rossi, B. (2006) "Are exchange rates really random walks? Some Evidence Point to Parameter Instability" Macroeconomic Dynamics, (forthcoming) 10.,1

Rossi, B. (2005) "Optical Tests for Nested Model Selection with Underlying Parameter Instability" *Econometric Theory*, 21(5).

Sarno, L., and M. P. Taylor, The Economics of Exchange Rates. Cambridge and New York: Cambridge University Press, 2002.

Taylor, M. P. (1995), "The Economics of the Exchange Rates", *Journal of Economic Literature*, 33, 12-47.

INSTRUCTIONS TO AUTHORS:

- The Arthaniti considers manuscripts for publication on the understanding that the submitted paper represents an original work not previously published, or being considered for publication elsewhere. The Journal has the copyright to all its published articles.

- Manuscripts accepted for review are blind refereed by two or more scholars. In this context, the authors are advised to write papers in a way to keep their identity from referees. The first page of the text should not reveal the name(s) of the author(s).

- Three hard copies of the manuscript along with an abstract in electronic form (not exceeding 100 words) should be sent to Dr. Kanailal Basak, Editorial Secretary, Arthaniti, Dept. of Economics, University of Calcutta, 56A, B. T. Road, Kolkata 700 050, India.

- Although there is no submission fee or page charge, each manuscript should be accompanied by a processing fee, which is Rs. 100 for authors within India and US $ 10 for authors outside India. Demand drafts must be made payable to "ARTHANITI".

- Manuscripts should be in English and between 4,000 and 8,000 words (15-30 pages). They should be double-spaced and typed on standard letter-size (11 x 8.5) paper.

- The cover page should include the author's name, institutional affiliation, and the title of the paper. Please include the contact person's full address, phone, fax numbers and e-mail address. Include the name(s) of any sponsor(s) of the research contained in the paper, along with grant number(s). Supply an abstract of up to approximately 100 words for all articles. It should also contain at least one classification code according to the JEL Classification and five keywords.

- Only essential mathematical notations should be used. All equations and statistical formulae should be neatly typed and numbered consecutively throughout the text in Arabic numerals.

- Footnotes should be kept to a minimum and numbered consecutively throughout the text with superscript Arabic numerals

- References should be quoted in the text as name and year within brackets and listed at the end of the paper alphabetically. Where reference is made to more than one work by the same author published in the same year, identify each citation in the text as follows: (Johansen, 1988a), (Johansen, 1988b). Where references include three or more authors the form (Johansen et al., 1988) should be used in the text. All references must be complete and accurate. Online citations should include date of access. If necessary, cite unpublished or personal work in the text but do not include it in the reference list. References should be listed in the following style: Johansen, S. (1988): Statistical analysis of cointegrating vectors, *Journal of Economic Dynamics and Control* 12, 231-254. World Bank. (1994) *Adjustment in Africa: Reforms, Results and the Road Ahead.* Oxford University Press, New York.

- Tables should be numbered consecutively in the text in Arabic numerals and printed on separate pages.

- Illustrations. Supply each illustration on a separate sheet, with the lead author's name and the figure number, with the top of the figure indicated, on the reverse. Supply original photographs; photocopies or previously printed material will not be used. Line artwork must be high-quality laser output (not photocopies). Tints are not acceptable; lettering must be of a reasonable size that would still be clearly legible upon reduction, and consistent within each figure and set of figures. Supply artwork at the intended size for printing.

- Final version of the papers must be submitted on floppy diskette in MS Word 2000 / Word Perfect 6.0 along with two hard copies.

- Final proofs will also usually be sent to the author for checking. This stage is to be used only to correct errors that may have been introduced during the production process. Prompt return of the corrected proofs, preferably within a week of receipt, will minimize the risk of the paper being held over to a later issue.

- A copy of the journal will be provided, free of charge, to each author.